

Initial View on 2021 Census Output Content Design

Response to consultation

Annex

Contents

Introduction	2
Evaluation criteria	4
Topics and geographies	6
Flexible dissemination system	9
Disseminating high volumes of aggregate data	21
Use of administrative data to improve and supplement outputs	26
Microdata	39
Origin-destination	44
Geography	53
Interpreting 2021 Census estimates: supporting information	62
Interpreting 2021 Census estimates: wider analytical uses of census data	68
Comparability of 2021 Census estimates: UK harmonisation and UK statistics	71
International comparability of 2021 Census results	78
Next steps	81

Introduction

The Office for National Statistics (ONS) invited views on our “Initial View on 2021 Census Output Content Design” from 28 February to 23 May 2018. This consultation outlined our initial proposed design of 2021 Census outputs and the dissemination channels for England and Wales.

Our vision is for flexible, timely and accessible census outputs. The main benefits of the census will only be realised by users having access to high-quality statistics, released free at the point of use and as speedily as possible.

Our initial view was based on:

- the evaluation of the design and delivery of 2011 Census output content
- our current understanding of user need and knowledge of potential technological and methodological solutions
- our current understanding of costs involved to develop and implement possible outputs and dissemination system
- our aspirations for using administrative data to improve and supplement outputs

Our consultation covered all aspects of 2021 Census outputs, including our plans for a flexible dissemination system to access the majority of census data and the approach for statistical disclosure control (which safeguards confidentiality). We also set out our plans for specialist products (including microdata and origin-destination products), geography (grids and parishes) and our aspirations for using administrative data to improve and supplement outputs.

We received 196 responses to the consultation. Of these, 130 were from organisations and 66 were from individuals. We’d like to thank everyone who gave us their views by responding to the consultation.

This report contains our response to the user feedback we received in the public consultation. We anticipate the information provided in this response will be of most interest to our more frequent users of census data.

We’ve attributed the names of respondents to the information they provided in our report, where the respondent has provided consent for us to do so.

As previously, there will be separate censuses in Scotland and Northern Ireland. We’re working closely with the National Records of Scotland and Northern Ireland Statistics and Research Agency to ensure UK comparability of results, where possible. NRS published a summary of responses to its 2021 Census Outputs Strategy in June 2018⁴. NISRA will be consulting with users in due course.

⁴ http://www.scotlandscensus.gov.uk/documents/census2021/Scotlands_Census_2021_Outputs_Strategy_Stakeholder_event_summary.pdf

Figure 1: Number of organisational responses by sector

Sector	Count	Percentage
Local authority	105	53.6
Academic/research	19	9.7
Charity and voluntary	16	8.2
Other	15	7.7
Commercial	13	6.6
Public body (for example health, transport, emergency services)	12	6.1
Government department	10	5.1
Genealogist/family historian	4	2.0
Housing	2	1.0
TOTAL	196	

2021 Census outputs roadshows

The consultation was supported by five regional roadshows held during April and May 2018 in Cardiff, Newcastle, Manchester, Birmingham and London. At the roadshows, we updated users on our initial proposed design of 2021 Census outputs, as outlined in our consultation. We also provided attendees with an opportunity to try a prototype of the flexible dissemination system.

The five roadshows were well attended by over 200 delegates in total, from various sectors. Overall, we received positive feedback on our approach for 2021 Census outputs, as well as useful and constructive comments about the flexible dissemination system design and user interface. We'll use this feedback alongside the evidence provided in this consultation and other user research to inform future iterations of the flexible dissemination system.

Evaluation criteria

The overall design of 2021 Census output content will be evaluated against the criteria in the following table. These criteria were agreed by the UK National Statistician when defining the scope of the 2021 Census outputs and dissemination project and are based on user feedback from the 2011 Census.

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

In drafting our response, we've assessed responses provided against the criteria of accessibility, flexibility, timeliness, user acceptability and public confidence. We identified where users had provided positive or negative comments relating to our intended design. We've also considered any other comments users made about our design that do not fit neatly into the above categories.

We also looked at the type of work being undertaken by users when considering the responses. This was based on the ONS Approved Researcher Scheme definition of research being undertaken to “serve the public good”⁵.

Types of work could include using census data for:

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

Where we have used percentages in the analysis within our response, we’ve used the number of responses to each individual question to calculate these percentages.

⁵ <https://www.ons.gov.uk/aboutus/whatwedo/statistics/requestingstatistics/approvedresearcherscheme>

Topics and geographies

We asked respondents to indicate their main topics of interest and the geographies they used the most after the 2011 Census. We'll use this information to help us plan our publication schedule. We'll consider it alongside:

- information gathered in the "Initial View on content for 2021 Census for England and Wales" consultation held in 2015
- download statistics for the 2011 Census from NOMIS and the ONS website
- our plans for census data processing
- aspirations to use administrative data

Topics

We wanted to know what the main topics of interest were for users following the 2011 Census.

Number of responses to question	196
Key findings	Age, ethnicity, sex, health, household composition and income were the most important topics

Geographies

We wanted to know which geographies have been used following the 2011 Census.

Number of responses to question	190
Key findings	Local Authorities, Wards, and Lower Layer Super Output Areas (LSOAs) were the most used geographies
	National Parks, other geography types, and GEOSTAT grids were overall ranked as the least-used geographies

This information is useful to understand the geographies that have been used. We'll consider this information alongside download statistics from the Open Geography portal and information from other public engagement. We're also aware users may indirectly utilise our geographies, for example the built-up area geography is the basis for rural-urban statistics.

Figure 2: Responses to the question "In 2011, which geographies did you use?"

MOST IMPORTANT	1	Local Authorities
	2	Wards
	3	Lower Layer Super Output Areas (LSOAs)
	4	Output Areas
	5	Counties/upper-tier local authorities
	6	Regions
	7	Parishes
	8	Middle Layer Super Output Areas (MSOAs)
	9	Countries
	10	Postcodes*
	11	Travel to Work Areas (TTWAs)
	12	Parliamentary Constituencies
	13	Workplace Zones
	14	Local Enterprise Partnerships
	15	Built-up Areas (including sub-divisions)
	16	Nomenclature of Territorial Units for Statistics (NUTS)
	17	National Parks
	18	Other
LEAST IMPORTANT	19	GEOSTAT grid

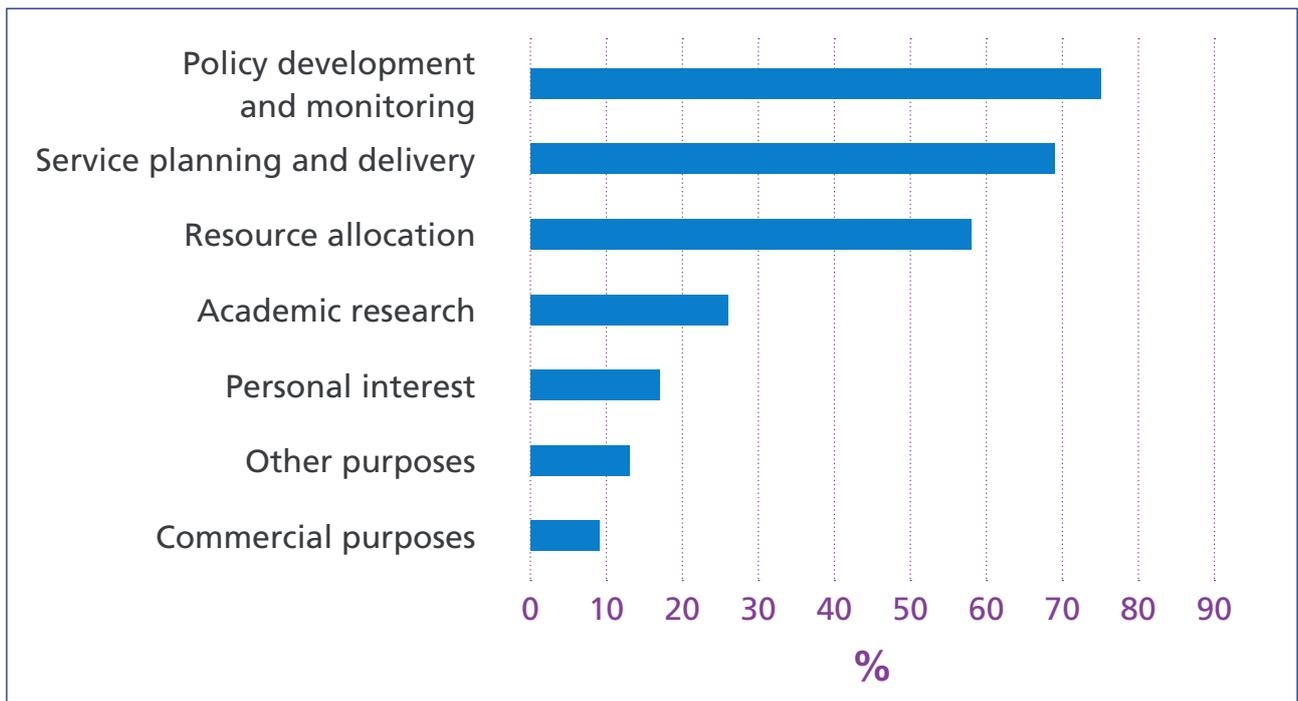
* 2011 Census Postcodes

Use of census data

We were also interested to understand how 2011 Census data had been used.

Number of responses to question	195
Key findings	Most frequent uses were for “policy development and monitoring” (75%) and “service planning and delivery” (69%)
	Lowest frequency of use was “other purposes” (13%) and “commercial purposes” (9%)

Figure 3: Responses to question “In 2011, what did you use census data for?”



Flexible dissemination system

For 2021, we're planning to disseminate census data through a flexible dissemination system. This will allow users to create their own datasets by selecting the geography, population base and variables they require. It'll mean users have more timely access to the data, compared to when they were able to access data from the 2011 Census. It'll also mean that census data are more accessible, as the majority of the data are accessed through one website. This approach will also utilise a statistical disclosure control methodology that uses targeted record swapping, a light-touch level of perturbation (known as the cell-key method) and automated table design checks. More information about this approach can be found in the consultation document⁶.

We wanted to establish the appetite amongst users for a flexible dissemination system and the statistical disclosure control methodology, and understand the impact this would have on their work.

Overall, 195 users responded to at least one question in this section.

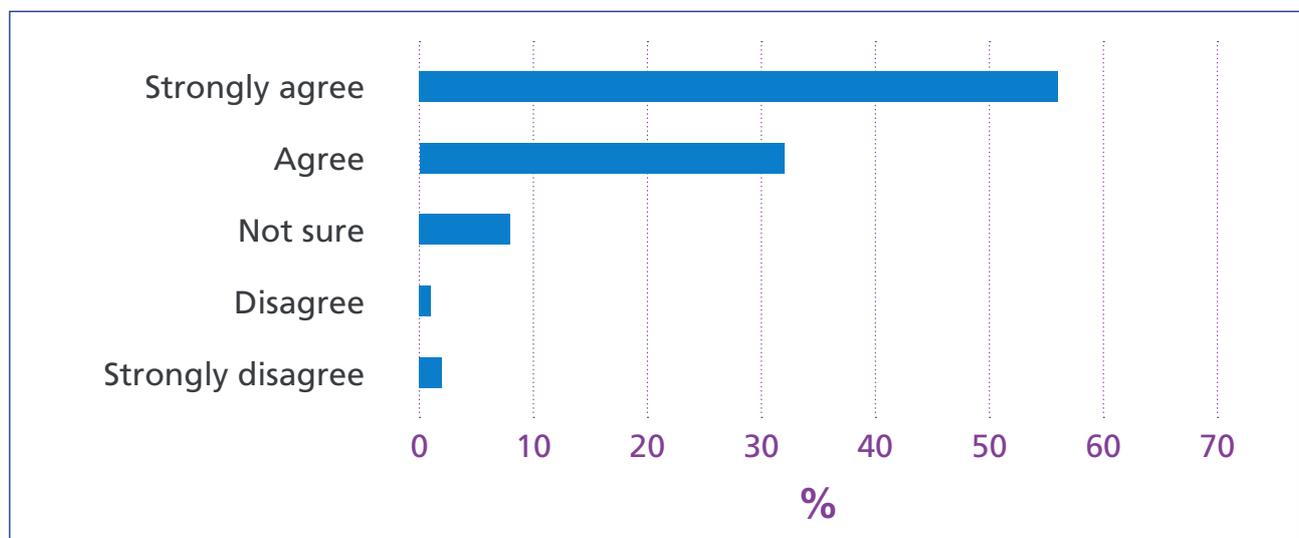
Number of responses to question	195
Key findings	88% either agreed or strongly agreed that they'd prefer to access data via an online flexible dissemination system sooner after census day, rather than wait for us to develop tables
	3% either disagreed or strongly disagreed
	9% were not sure

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

We asked if users would prefer to access data using an online flexible dissemination system sooner after census day than if they were to wait for us to develop tables.

Number of responses to question	195
Key findings	88% either agreed or strongly agreed that they'd prefer to access data via an online flexible dissemination system sooner after census day, rather than wait for us to develop tables
	3% either disagreed or strongly disagreed
	9% were not sure

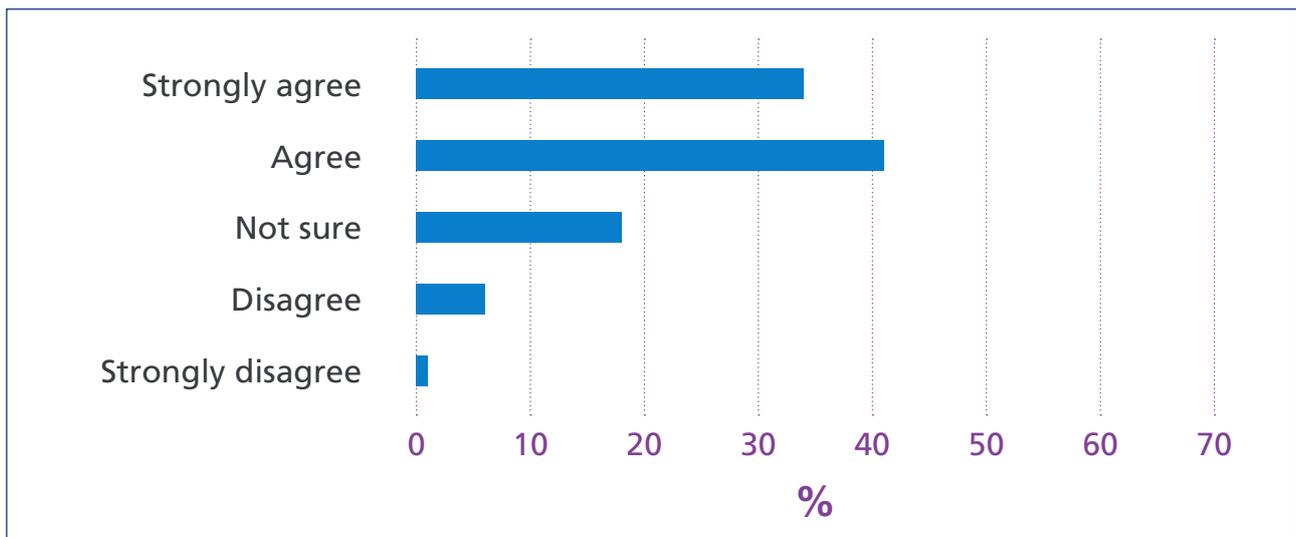
Figure 4: Responses to question "I would prefer to access data via an online flexible dissemination system sooner after census day than wait for tables to be developed by ONS".



We wanted to know whether users thought defining their own tables is more useful than using predefined tables.

Number of responses to question	194
Key findings	75% of respondents either agreed or strongly agreed that the ability to define tables was more useful than using predefined tables
	7% either disagreed or strongly disagreed
	18% were not sure

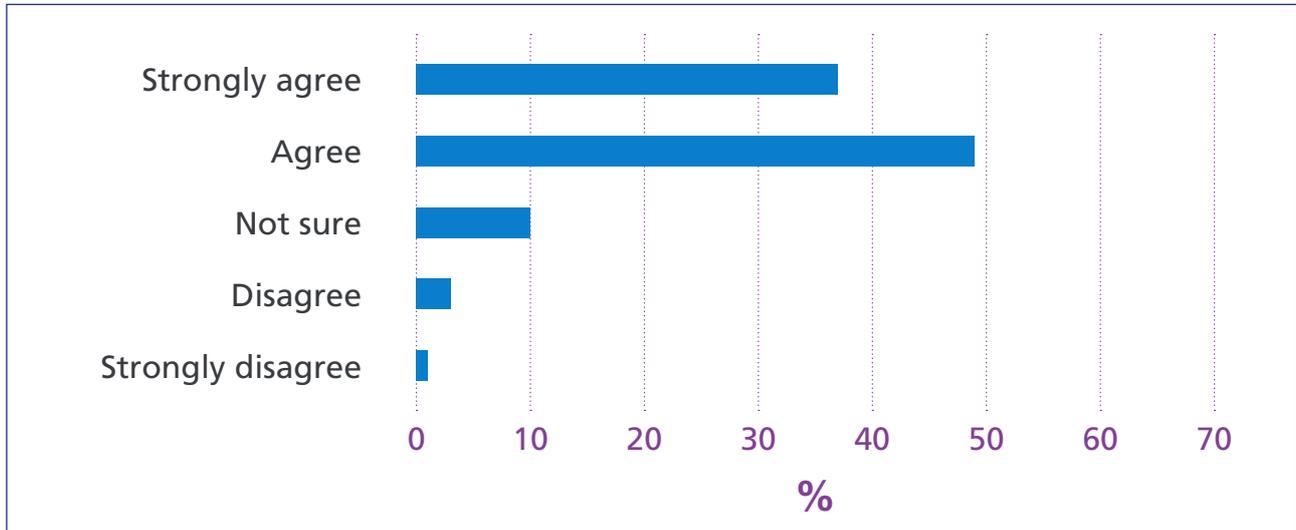
Figure 5: Responses to question “Being able to define my own tables is more useful to me than predefined tables”.



We wanted to identify whether users would find it useful to be provided with a small number of univariate tables alongside those generated through the flexible dissemination system.

Number of responses to question	193
Key findings	86% of respondents either agreed or strongly agreed that they'd find it useful to be provided with a small number of univariate tables alongside those they can generate through the flexible dissemination system
	4% either disagreed or strongly disagreed
	10% were not sure

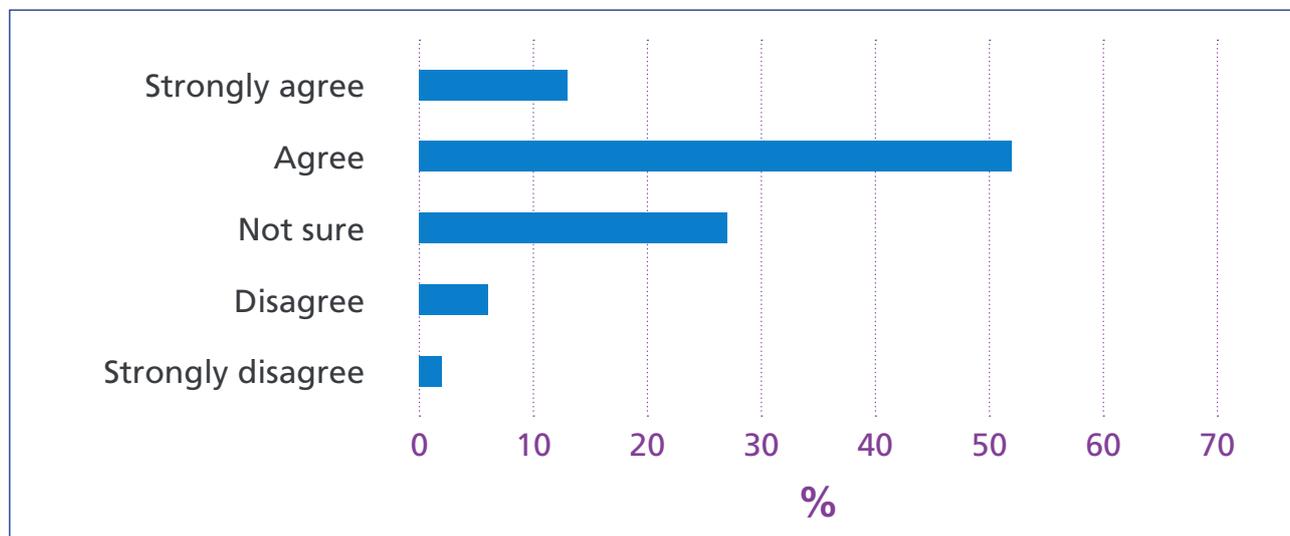
Figure 6: Responses to question “I would find it useful to be provided with a small number of univariate tables alongside those I can generate through the flexible dissemination system”.



We also wanted to understand the impact of the proposed statistical disclosure control methodology on users and whether it'll be acceptable to use tables where there are small inconsistencies in output table totals due to the perturbation caused by the cell-key method.

Number of responses to question	192
Key findings	65% of respondents either agreed or strongly agreed that they'd find it acceptable to use tables with small inconsistencies in output table totals due to the perturbation caused by the cell-key method
	8% either disagreed or strongly disagreed
	27% were not sure

Figure 7: Responses to question "I would find it acceptable to use tables where there are small inconsistencies in output table totals due to the perturbation caused by the cell-key method".



We also received 66 responses from users providing more detailed feedback on our proposed flexible dissemination system.

Timely release of census data

The earlier, more timely access to 2021 Census data compared to 2011 Census data and not waiting for us to produce tables was considered beneficial by users. This is especially when they need data on specific population groups.

Tower Hamlets Council: *“Potentially, this could provide valuable data on small population groups (ethnic groups and individual countries of birth) avoiding lengthy waits for commissioned tables - so is a very welcome development.”*

Cornwall Council: *“...write-in options should be included within the FDS [flexible dissemination system] which will then enhance the ability to determine the cross tabulations needed rather than relying on pre-determined tables. Write-in options should also be made available at the same time as the main release.”*

However, we received one comment stating that it would be more valuable to users for us to consider the order in which we released data.

The Market Research Society, and the Market Research Society’s Census and Geodemographics Group and Data Analysts User Group (DUG): *“Can the priority of outputs being released be ordered by user demand/usefulness? For example, can outputs related to age and sex be released first? In 2011, age by sex at output area level was one of the last outputs to be released. It really should have been one of the first because it is fundamental to a vast range of users.”*

There were also concerns raised about our ability to deliver the flexible dissemination system. Plus, there was a request for any system to be available before the release of data for users to test and become familiar with it prior to the release of census data.

Caerphilly County Borough Council: *“I have some concerns regarding the flexible dissemination system - something of this nature was promised in 2011 and nothing happened. I can understand the reasoning behind providing users with the flexibility to design their own tables and the reduced workload for ONS in terms of providing ready-made tables. However, this system would need to be proven to work well in advance of the Census results being made available. I would suggest expert users of Census data are allowed access to the tool for test purposes (such as beta versions) as early as possible, to ensure that it meets their needs.”*

London Borough of Camden: *“We are not convinced that the flexible dissemination system will provide the full range of statistics at the differing levels of geography that we require. Previous census products that were meant to provide such utility have failed to deliver and, therefore, there is a concern that throwing the eggs into*

this one basket is a huge risk. A standard table-set would ensure that users could rely on a guaranteed set of outputs, even if this takes longer to complete. Starting off from set tables is a good way to carry out a sense check and spot if something is going wrong – you're also more likely to pay attention to notes relating to the table population and the individual variables."

Functionality of the flexible dissemination system

The proposed functionality of the flexible dissemination system, with the ability for users to define their own datasets, was also welcomed.

Cornwall Council: *"The flexible dissemination system is very much supported, providing a great deal of flexibility alongside timely dissemination."*

Users welcomed our proposal to improve the accessibility of census data and ensure most data are available through a single point of access.

Basingstoke and Deane Borough Council: *"The proposal to host the majority of census data from one location is welcomed, especially as the publication of the parish data separately on Neighbourhood Statistics caused confusion when the 2011 data was released."*

There was also an indication that the format of data produced by the flexible dissemination system is important for users to provide compatibility with their own software and systems.

A county council: *"Programmes such as R do require a level of standardised reporting and a rigid template of the csv output from the dissemination tool. Therefore, a predictable and consistent excel output (preferably unformatted csv) would be desired. Templates/Layouts of the proposed output before the census data is released (could be filled with dummy variables), such that scripts for R could be created in advance of the census release would also be of value to us."*

The Market Research Society and the Market Research Society's Census and Geodemographics Group and Data Analysts User Group (DUG): *"Table output export options must include MS Excel and csv file formats."*

We also had several comments that the existing NOMIS website is heavily accessed for 2011 Census data and users would like to see the flexible dissemination system providing at least the same functionality already provided by NOMIS.

Central Bedfordshire Council: *"NOMIS was used extensively for the 2011 Census results and is still used now, so any new system needs to have at least the functionality of NOMIS, both in terms of site design and user support"*.

Rotherham Metropolitan Borough Council: *“Please build on NOMIS as an effective way to access Census data and also the table building capability of StatXplore for DWP data which could be duplicated for Census data.”*

We also received feedback that the design of the flexible dissemination system should be accessible for different types of users, not just individuals and organisations who are familiar with census data and use it on a regular basis.

Cornwall Council: *“It should also be recognised that there is a huge variation in user knowledge and so consideration should be given to the inclusion of a search facility based on postcode or place name, such as that on NOMIS. This would support those not familiar with Census Geographies and coding to access the data needed.”*

We also had feedback from one respondent who emphasised the importance of providing a version of the flexible dissemination system in the Welsh language.

Gwynedd Council: *“Important that any dissemination system is fully bilingual (Welsh/English) with Welsh-language content properly tested/proof-read beforehand.”*

The proposed flexible dissemination system will allow users to build their datasets. However, a concern was raised that if census data are revised following publication, then users will be required to rerun their queries to recreate their dataset. The creation of tables by the user was also considered to be an onerous process.

Edmund Fallon: *“My only concern would be that if we get all the data first and do not wait for ONS to create them, as they have done in the past, then we may run the risk of inconsistencies in analyses and ongoing revisions to analyses which have been self-created. There is some comfort in knowing that the ONS have ratified all tables before their release.”*

London Borough of Camden: *“Defining tables for ourselves will create work at a time when we have fewer people and less time. With a lot of user defined tables I think we might find ourselves comparing apples and pears quite often where different users pick different variables as indicators of the same thing.”*

We also had a concern about the performance of the flexible dissemination system and its ability to cope with a large volume of users.

The Market Research Society and the Market Research Society’s Census and Geodemographics Group: *“This system [needs to] cope with the very heavy demand that will come from a vast array of users all wanting these data at the same time.”*

Pre-determined tables

In our consultation document, we mentioned that we're planning to produce some pre-determined tables alongside the flexible dissemination system. A number of respondents emphasised it was emphasised by a number of respondents that publication of some pre-defined tables is very useful and important to local authorities.

Cornwall Council: *"The provision of predefined tables is however also important as some of the key statistics tables are staples for many local authorities and would reduce duplication of efforts. This may also provide an alternative means of accessing data, albeit limited, should high level traffic on the FDS be such in the initial release phase that access is problematic."*

Hertfordshire County Council: *"Having a set of univariate tables gives us an immediate overview of the results at local authority level without having to try and guess which variables we would need to choose from a system. These were predetermined for 2011 based on the layouts from 2001, also giving comparability."*

Comparability with 2011

Users also noted it's important there's functionality available that allows comparability of 2021 Census data with 2011 Census data.

Greater Manchester Combined Authority: *"We welcome a flexible approach to data dissemination. However, we would also have a need to compare easily with the 2011 and 2001 Census, and in this respect metadata accuracy is also important."*

A county council: *"It would be helpful to have 2011 data available - perhaps a comparator tool so data can be manipulated in the same place rather than having to go back to Nomis and get the data as it is currently."*

Statistical Disclosure Control Methodology

Our proposed statistical disclosure control methodology will produce some small inconsistencies between table totals due to the cell-key perturbation method. Overall, most of the feedback about our proposal to utilise a flexible dissemination system and statistical disclosure control methodology was positive. Users would be satisfied with small inconsistencies in the totals, providing that guidance was provided for users.

London Borough of Camden: *"I would find it acceptable to use tables where there are small inconsistencies in output table totals due to the perturbation caused by the cell-key method. We have been here before, but it's not an easy sell to the less savvy user."*

City of London Corporation: *“However, personally I think small inconsistencies are fine. From an end user perspective, my simple approach would be rounding data to at least 10 for small areas and probably at least 100 for larger areas. In simple terms, it would have minimal impact. The key in my opinion is provide guidance on using the data and quoting in analytical reports e.g. using a set of rounding guidelines.”*

Stevenage Borough Council: *“The flexible dissemination tool will enable a vast number of different tables to be produced. Perturbation will mean two similar tables may provide different figures. It will therefore be essential that any downloaded files are clearly marked with the precise table configuration, and/or some kind of reference or URL that will enable the data to be referenced effectively and downloaded by others.”*

However, some users did express apprehensions at our overall approach and were particularly concerned about the effect of applying the cell-key perturbation method and the impact on table totals.

Data Analysts User Group (DUG): *“We do not want a more stringent SDC method than the one used for the 2011 Census as this would have a detrimental effect on outputs.”*

Basingstoke and Deane Borough Council: *“We are against the use of a flexible dissemination system because of the inconsistencies in output table totals caused by perturbation. This will cause confusion and distrust of the data amongst users, particularly decision makers, as well as creating difficulty in every day analysis and for these reasons we do not think it is acceptable. From memory, this was exactly the problem that users had with the 2001 Census data and the very reason that 2011 Census data was developed to use record swapping which allowed consistency across all tables.”*

Our response - Flexible dissemination system:

We welcome the feedback users have provided us about our plans to utilise a flexible dissemination system to disseminate 2021 Census outputs. We're pleased with the mainly positive response users have provided, but we recognise this approach may not satisfy all users. We'll continue to work with users to understand their needs and develop the final design of census outputs to meet these needs as far as possible.

We're currently developing the statistical disclosure control methodology and seeking assurance of our methods from a wide range of experts. The UK National Statistician is supporting the development of the statistical disclosure control approach.

A aim for our outputs is to ensure users will be able to access at least the same level of detail in variables, geographies and population bases as in 2011. In 2011, the statistical disclosure control methodology meant that, if one geographic area failed a disclosure check, then data would not be available for any area. The new methodology proposes that data will be available for any areas that pass disclosure checks, regardless of whether another area fails. This will lead to more data being publicly available.

The cell-key method will enable more timely and flexible outputs. It will also protect against disclosure by differencing in cases where every geographic area except one could be released from a larger geographic area. We're continuing to investigate the size of inconsistencies between table totals caused by the cell-key perturbation, but we anticipate any differences are likely to be small. The level of perturbation will be set to ensure inconsistencies are minimised whilst protecting against disclosure. We'll provide guidance to users to explain the statistical disclosure control methodology and how to use the data.

We have identified there may be some census data which will not be available via the flexible dissemination system, but will still have the same statistical disclosure control methodology applied, for example data on small population groups derived from topics that are new to the 2021 Census. We are exploring how to provide this data in an alternative format, for example in a pre-determined table.

We will also continue to provide a service for creating bespoke or commissioned outputs that have not been included in standard output products following the 2021 Census. We will apply the same statistical disclosure control methodology as all other standard tabular outputs.

We've worked with an external company to build a prototype flexible dissemination system and to demonstrate its feasibility. We're now seeking a solution that can be built within the ONS IT architecture and developing the user interface on the ONS website. The feedback we've received from both the consultation and roadshows indicates that users are familiar with the functionality

offered on the NOMIS and now closed Neighbourhood Statistics (NESS) websites. We want to continue our engagement with users to further understand the specific functionality and preferred design elements, and aim to incorporate these into the flexible dissemination system design. We're also working to ensure the flexible dissemination system provides the option to access data in Welsh, where applicable.

We're aiming to have a fully functioning version of the flexible dissemination system on the ONS website before the release of 2021 Census estimates. To increase the timeliness of 2021 Census data, we would like users to become familiar with the online system and have the relevant programmes embedded in their systems so they can access the ONS Application Programming Interface (API). We'll also be stress-testing the flexible dissemination system before its launch to ensure it can withstand the expected volume of users.

Estimates derived from the 2021 Census estimates will be fully quality assured using a robust methodology before they're published on the ONS website. For outputs based upon integrated administrative data, we'll ensure the accompanying metadata summarise the source, definitions and methodology used to derive the estimates. In the event of a correction being required to census data after publication, we'd follow the revisions policy⁷ which will outline the process for amending data and notifying users of this change.

It's our intention to utilise the benefits of the flexible dissemination system beyond outputs from the 2021 Census. Its application will allow users to access wider ONS data, including outputs based on administrative data.

⁷ <https://www.ons.gov.uk/methodology/methodologytopicsandstatisticalconcepts/revisions>

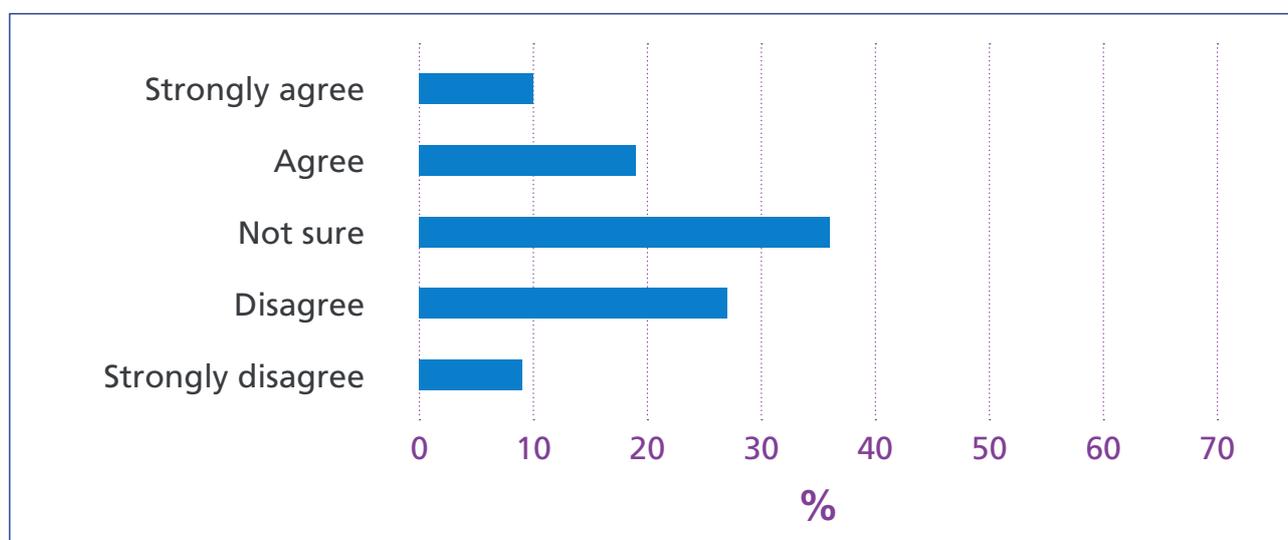
Disseminating high volumes of aggregate data

We recognised the requirement from some users for high volumes of aggregate data (especially by those that add value and redistribute the data) and for queries to be submitted by systems outside of ONS.

We asked users if they need a high proportion of aggregate data to install into their own systems, for example bulk data.

Number of responses to question	192
Key findings	29% of respondents either agreed or strongly agreed they needed a very high proportion of aggregate data to install in their own systems
	36% of respondents either disagreed or strongly disagreed
	36% were not sure

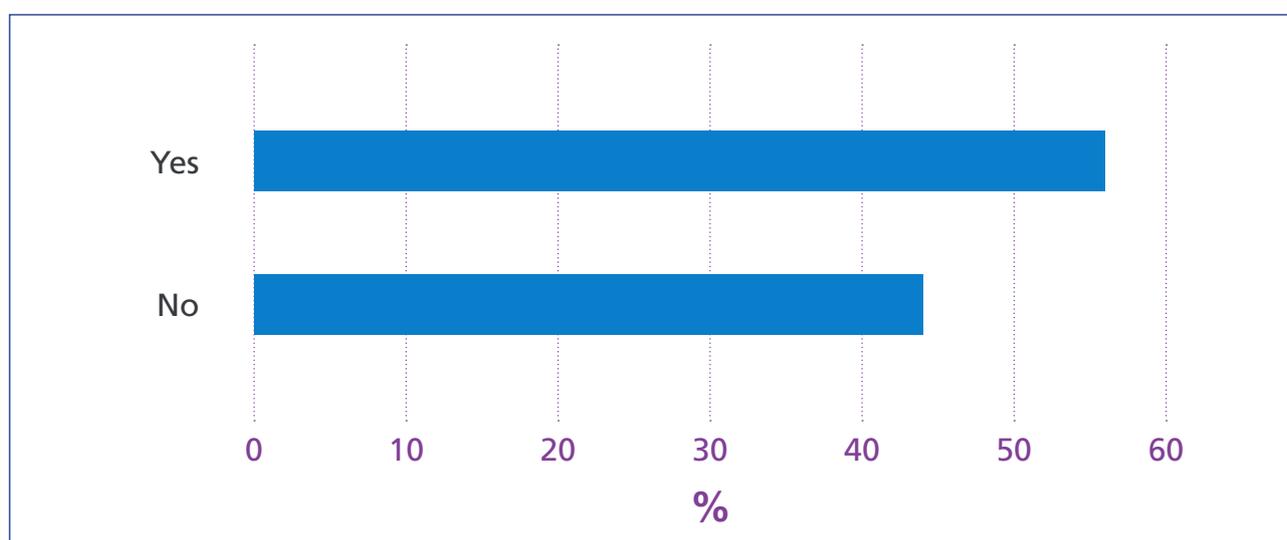
Figure 8: Responses to question “I need a very high proportion of aggregate data to install in my own systems (for example bulk data)”.



We also wanted to identify whether users would be able to utilise an Application Programming Interface (API) to access census data.

Number of responses to question	172
Key findings	56% of respondents stated that their systems would be able to utilise an API to access census data
	44% said their systems would not be able to do so

Figure 9: Responses to question “Would your systems be able to utilise an API to access census data?”



We asked how users would expect to make use of data via an Application Programming Interface (API). An API enables separate web systems to communicate with each other in an automated way. We received 94 responses to this question.

Users told us there was a need for large volumes of data.

GEOLYTIX: “... a bulk download of the Key Statistics tables would still save us and I suspect several others a painful processing task.”

Data Analysts User Group (DUG): “A Bulk Data option is key to us. A lot of our users want to be able to download all of relevant tables at output area level in one go. Our users will be using the vast majority of outputs so we would not want to repeatedly have to select each table or variable in turn. We want as rich a set of outputs in terms of depth and breadth as possible.”

Users from national government, local authorities, the commercial sector and academia clearly indicated a demand for an API to access census data, and they already had the functionality in place to access and use an API.

Users mentioned they would use an API to bring census data into their systems to undertake analysis, or use it to combine census data with data from other sources in real time.

Dr David Owen, University of Warwick: *“Would use the API to draw data into programs written in R and/or Python to calculate measures based on Census data for large numbers of small areas.”*

Gwynedd Council: *“We’d like to explore the possibilities for combining internally sourced data (client, service provision etc) with external data (Census and others) to produce information and performance ‘dashboards’ which are then updated in real time in line with the ‘open data’ concept.”*

Functionality of an API

The functionality of an API would allow users more timely access to data and save them time and resources when downloading and accessing the data.

Cornwall Council: *“An API would be used to feed a dataset that would be used within our Business Intelligence System. This would then be used within interactive dashboards and combined with other related datasets that are then used across the organisation for service planning and resource allocation purposes. We may look to generate specific reports using the data from the API which would allow us to publish Cornwall specific 2021 Census data on our website in a timely manner.”*

East Riding of Yorkshire Council: *“We have a cloud-based hosted website that our supplier obtains data through APIs to populate (freeing up officer time to actually use the data, rather than spending the time to upload it).”*

A county council: *“We use this [API] for downloading key economic data as it saves some time compared to building upload sheets in excel.”*

Stefan Noble: *“We plan to extract all of the census data tables (including cross tabular data) at Output Area and import into our internal databases to aggregate to the bespoke geographies that we work with. This is a very large job if undertaken manually so being able to make use of an API will speed the extraction and import process.”*

Benefits of an API

The benefits of using an API would also increase accessibility to the census data for the analysts and for their internal and external stakeholders.

Salford City Council: *"We are also exploring the option of increasing openness by allowing residents to interact with local data directly, including census data through data visualisation software such as Tableau. APIs would streamline and improve this activity."*

House of Commons Library: *"An online API would make it easier to automate and standardise the production of these reports and analyses. It would also allow us to use data in online applications (we have previously used Census data from the Nomis API for mapping ethnicity by Parliamentary constituency, for example)."*

Pembrokeshire Coast National Park Authority: *"To create/embed a 'live' National Park Profile on our website."*

Greater Manchester Combined Authority: *"We currently access a variety of information via APIs to support our mapping tools on MappingGM. Access to a system that provides Census information in an API will allow us to visualise data at a local level for stakeholders across GM, and allow comparison with previous years Censuses. Additionally, a system like this may also allow more information to be easily mapped through our own internal GIS systems, and those of our partners."*

However, we're aware of the importance of the timeliness of the delivery of an API that's fit for purpose.

Greater London Authority (also incorporates input from SASPAC): *"We worked closely with ONS to deliver the 2011 Census data to our users (which include the GLA and the 33 London Boroughs) and were members of the Output WG and Bulk Outputs WG... Although much time/energy was invested in developing an API for the 2011 data, this failed to deliver anything meaningful to users. The outcome was a very limited set of data, delivered after the event via an API that was of no use to SASPAC, and imagine, anyone else. For the API to be of use, it needs to contain all the data/geography for which outputs are being produced and delivered at the same time as data through any other ONS system. It will also need to be available for testing well-ahead of the first data release, ideally serving up 2011 data, and allow us to develop a system that can consume the API. Crucially, the system needs to be mature enough to handle the 2021 outputs (structure/metadata/code lists/geography lists) well before the data arrives."*

Overall, the responses indicated there was a willingness and user need to use an API. However, they also indicated that, at present, some users do not have the level of expertise, or their systems do not have the functionality for, an API.

Central Bedfordshire Council: *“It is possible we would use it with Public Health data, but we will need to develop our technical knowledge.”*

Hertfordshire County Council: *“Obviously an API is a great way of disseminating but the level of expertise needed to use one is above our level.”*

Our response - Disseminating high volumes of aggregate data:

The responses indicate there’s a clear user demand for an API to access census data. We’ve many users who already use APIs in their work and would value this functionality to reduce the time and money required to import data into their systems.

Our aim is to work closely with users to understand their requirements for the delivery of high volumes of aggregate data and develop an API that meets their needs.

We’ve already started development of an API⁸, so we’ll build on this work to develop an API to access 2021 Census data. The range of data available through the API would be same as provided through the flexible dissemination system and user interface on the ONS website.

We’re aiming to produce an API in advance of the dissemination of census data to allow the timely delivery of data and give users the chance to embed this in their systems.

We’ll also provide guidance to enable users to understand and fully utilise the API and promote its availability.

⁸ <https://developer.ons.gov.uk/office-for-national-statistics-api/reference>

Use of administrative data to improve and supplement outputs

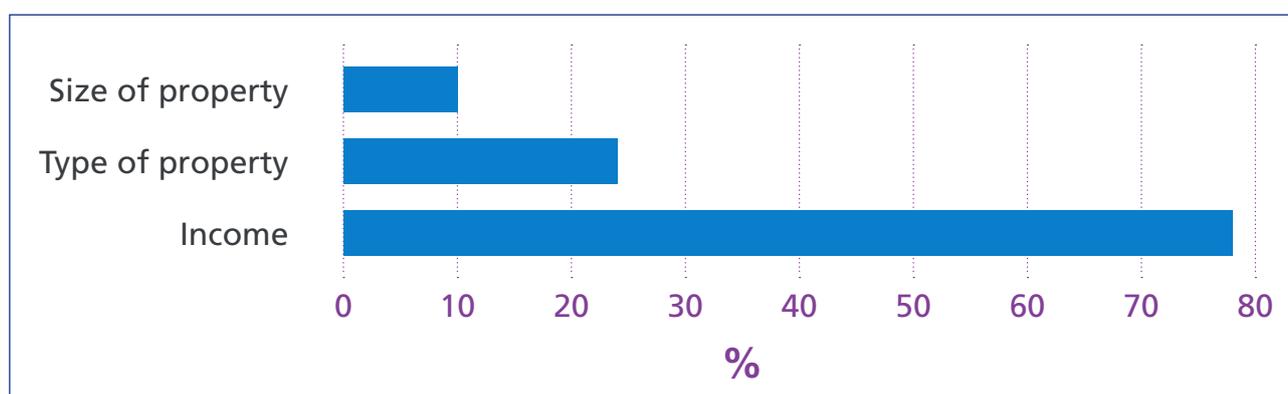
To complement the release of census data, we intend to create outputs where census data will be enhanced by integration with other data sources. We're currently exploring which administrative data sources could be used to produce these outputs.

We asked users to indicate which topics not currently collected on the census questionnaire they'd be most interested in receiving data about.

Number of responses to question	179
Key findings	78% stated they would be most interested in receiving data about income
	24% said they would be most interested in receiving data about type of property (a more detailed breakdown than available from the census)
	10% would be interested in receiving data about size of property

Although this question was not designed to be multiple response, these results have been presented to show where respondents did choose multiple response options. Respondents were only presented with these three options.

Figure 10: Responses to questions "Which topics, not currently collected on the census questionnaire, would you/your organisation be most interested in receiving data about?"



We wanted to know the impact on the users' work of having this additional data. Users provided comments on their need for data about income, a more detailed type of property, and size of property. There were 148 responses to the questions about using administrative data to produce supplementary outputs. Users indicated a clear user need for this information and highlighted the benefits it would bring to their work.

Income

Users provided us with several examples of where income estimates derived from administrative data would assist their analysis and service delivery, particularly in understanding social inequalities, deprivation, social exclusion, travel patterns and health in local authorities.

Leicester City Council: *"Income - Leicester has high levels of deprivation across the city. A large proportion of the population reside in income deprived households. The ability to analyse census categories with income will provide us with a wealth of information about some of our vulnerable communities."*

North Yorkshire County Council: *"It would improve our ability to target resources to reach low income households, increase our depth of understanding of poverty issues and allow pockets of hidden deprivation within largely affluent areas to be identified."*

Leicestershire County Council, Business Intelligence Service: *"This data would help us to understand our communities, alongside other economic data, and to help classify service users based on the income of their place of residence."*

Lewes District Council and Eastbourne Borough Council: *"It would provide a more holistic understanding of circumstances within the district/borough, enabling us to make more robust conclusions and to be more confident in policy direction and for planning service delivery and resource allocation, for example."*

Salford City Council: *"For affordable housing strategy, housing strategy and policy development we need calculation of affordability ratio at different geographies across the city, therefore access to some form of income data is essential. Additionally, this information is used to identify specific areas to target for intervention. For example, areas with concentration of low income families need targeted intervention to prevent fuel poverty."*

Department for Environment, Food and Rural Affairs (DEFRA): *"We might be able to get a fuller picture on farm household incomes rather than just incomes from farming."*

Cheshire East Council: *“The income data would help us to assess geographical variations in our residents’ standards of living and the extent and nature of their reliance on public services. For example, people on lower incomes have less money to spend on transport costs and so may depend more on public transport than other residents do. It would also help us assess geographical variations in housing affordability, which in turn helps to inform planning policy decisions about the provision and location of housing in general and of affordable housing in particular.”*

Shropshire Council: *“The data is required to understand the ability of households to afford different types of housing tenure in Shropshire at a sub-county level i.e. to calculate an affordability ratio (normally lower quartile house price to lower quartile household income).*

Affordability ratios are considered a key market signal in National Planning Policy Guidance to local authorities developing their Local Plan. They are used to;

- *Identify affordable housing need at sub-county level and to determine the future housebuilding requirement,*
- *Exam access to housing by type of tenure in order to inform policies influencing the type and tenure of housing required,*
- *Provide evidence to underpin negotiations with developers over the construction or contributions to affordable housing,*
- *Provide evidence to underpin decisions on planning applications,*
- *Provide area profiles of market town etc. to aid web user’s needs.”*

Opportunities for multivariate analysis using income data

There were many users who stated income data could be valuable when used alongside other census variables.

Home Office: *“Information on income would inform what is known about the socio-economic characteristics of migrants, and permit more detailed impact assessments for relevant policies.”*

Tower Hamlets Council: *“Ability to map income levels for different population groups would be invaluable. While we have access to various data about income from non-census sources, there are significant gaps, particularly at local authority level. For example, to enable us to map income levels by ethnic group or disability.”*

East Sussex County Council: *“It would be useful to get a better understanding of income by households for all ages. As a county with a relatively high proportion of households of pensionable age we would find it helpful to understand income distribution of the retired as well as the working age populations. This would inform service planning in Adult Social Care.”*

Irish in Britain: *“Housing and Income issues are of interest to us in the context of the equalities direction of our work so that any development/extension of data in both of these areas would be of relevance. The issue of Ethnicity is a key factor for us and it would be important for us that any new data in these areas could be cross tabulated with Ethnicity.”*

Humanists UK: *“Humanists UK is primarily interested in the religion question. Combining the administrative data on income would allow us to cross-tabulate religion and belief with a wider range of socio-economic factors for academic purpose, planning and resource allocation and for campaign planning.”*

Migration Observatory at the University of Oxford: *“Good income data (particularly on unearned income and self-employment earnings at individual and household information) are not currently available in the datasets that include migration-related variables, so having this information from Census would be valuable. We would use it for policy-focused research to understand the integration of different migrant groups in the UK, as well as to understand the economic impacts of migration.”*

Users also noted how they’re particularly interested in income data to help understand health outcomes.

A public body (health): *“Could connect to health data to see if there is any correlation between income and long-term conditions e.g. diabetes.”*

Public Health Wales Observatory: *“We would use information on income, size of property and type of property. This would give us the ability to analyse the impact of these topics on health which would be used by public health professionals aiming to improve population health.”*

Equality and Human Rights Commission: *“Income data linked to Census data would provide a means of comparing income between different groups as well as allowing associations between income and other topics to be explored, for example with health.”*

A local authority: *“Income indicators could help our intelligence team identify areas in Wirral where there are high levels of income inequality, this in turn will inform Public Health commissioners of where there are greater areas of need in Wirral.”*

Reducing the need for proxy data and improving accuracy

Several users stated how they currently use other data sources as a proxy for income data in their work, so would welcome a more robust measure of income. However, users mentioned how these proxy data sources have various limitations due to accuracy and geographic level.

Rotherham Metropolitan Borough Council: *“Useful to target low income areas rather than use proxy measures which are less accurate.”*

Braintree District Council: *“It would be useful information for most of the Council services and policy development. It would be useful to consider in the context of housing affordability, for example, which is critically important and largely determines the housing target for plans, but at the moment we only have data that uses ASHE sample survey data and that is volatile at LA district level, bearing in mind it is lower quartile earnings (i.e. a subset) that is looked at. It would be useful to have a comparator from the Census as a check, and this could give us workplace based as well as residence based comparisons.”*

Bristol City Council: *“Income data for small areas cross tabulated with other variables would be extremely valuable, even more so if it was possible to get fairly regular updates and track change over time. At the moment, we rely on the indices of deprivation for understanding deprivation across the city but without the added information on the types of people/households affected. Also, any deprivation index is quickly out-of-date - the CLG 2015 Indices of Deprivation is mainly based on data sources as at 2012/13. We are now in 2018 and won’t get any update until summer 2019 at the earliest! Even then it will be based mainly on 2017 data.”*

The Market Research Society and the Market Research Society’s Census and Geodemographics Group: *“For years our users have wanted an income based question or variable to be included in the Census. Using variables like number of cars, size of house, qualifications, occupation, qualifications have been used for years as proxy indicators of income, but they are not the same nor as accurate. Adding an income output or outputs would greatly enhance the Census and the opportunities available to our users for our work, particularly in building geodemographic classifications.”*

Reduced costs for data users

Users were interested in obtaining income data from ONS as they currently have to pay for this information from a commercial company or third-party provider.

Hull City Council: *“It would mean we would no longer have to purchase commercially available data on household income. At the moment, we use the Index of Multiple Deprivation to understand whether households are potentially affluent or deprived but we don’t get a sense of their income.*

Income would help us have better understanding of actual what the household earns - currently we have to buy this kind of modelled data from organisations such as Experian.”

East Riding of Yorkshire Council: *“Income data would be useful as currently to obtain this data locally, we have to pay for it commercially.”*

Caerphilly County Borough Council: *"We currently have no access to this sort of data. Whilst there are commercial sources such as CACI Paycheck these are often expensive, particularly in light of ever reducing budgets."*

Northumberland County Council: *"Having more data on income levels would mean we would not need to buy data on income from commercial companies."*

Shropshire Council: *"At the moment, the Council has to purchase Gross Household Income data from a commercial data supplier. The licence parameters are quite restrictive and make it difficult to share information outside of the Council."*

Definition of income variable

Users provided valuable comments about the need for income data. However, they also indicated how there are varying definitions of income, both individual or household. Also, definitions of income vary depending on the source of data used.

Greater London Authority (also incorporates input from SASPAC): *"Any income would be welcome, but need to think about what income measure(s) would be most useful: household income is more meaningful."*

London Borough of Hackney: *"We are aware through discussions at Network meetings with ONS colleagues that a number of matters relating to Income have not been resolved. For this to be robust and valuable data the definition of Income should be in the widest sense and includes sources of income not picked up through PAYE and self-assessment. This is challenging but the data will be of less use if it excludes types of income because it is more difficult to identify."*

London Borough of Camden: *"Income would be useful to cross-tabulate with other variables. It remains to be seen what definition of income would be used, which would affect its utility."*

Dr David Owen, University of Warwick: *"Thus far, other variables have been used as a proxy for income. However, income is a complex phenomenon, and a single 'income' variable may not add much information - e.g. it is necessary to know the sources of income for a household (work, benefits, etc.) and individual, and work pattern in order to properly interpret income."*

Joseph Rowntree Foundation: *"If we had income level it would transform how much use we would get out of the Census. Our overarching remit is to understand the root causes of social problems and solve poverty. Income would allow us to split the population into income quintiles and compare the bottom fifth with the top fifth and to calculate the poverty line (below 60% of the median income). With this in place we'll be able to look at all variables by whether an individual or a household is in poverty or not. It does however depend on how income is categorised as to whether we are able to use the calculation that will identify individuals and of households in poverty."*

Users also stated that if an administrative data source was utilised, there would need to be clear metadata describing the variable and its accuracy.

Hampshire County Council: *“We would be interested to understand the level of accuracy in these data (esp. income) and some simple accompanying notes that explain this so that we can then understand and explain this to others so that any new data can be used appropriately.”*

Justine Rego: *“The robustness of the enhanced data categories will need to be very clearly written out so that statisticians can easily explain the limitations of the data to managers and service planners.”*

Geographic level of income data

Users also told us how income data would be beneficial at a range of geographic levels, particularly below lower layer super output area (LSOA).

Gwynedd Council: *“Very difficult at the moment to obtain reliable income data, especially at lower geographical levels and covering all types of income. This information would help gain a better understanding of household circumstances/ needs and enable better targeting of resources and interventions.”*

Norfolk Constabulary and Suffolk Constabulary: *“Information on income for a range of geographies is of interest (the English Indices of Deprivation currently present this information at LSOA level only).”*

Professor Martin Phillips, University of Leicester: *“It would be extremely helpful if provided at a fine geographical scale - e.g. OA or LLDOA [LSOA]”*

Our response - Income:

Over the last three years, we've developed personal and household total income estimates for small geographic areas based on administrative sources. We'll continue to develop these census-type income data, so that they can be integrated with the data collected in the 2021 Census. This would enable multivariate analysis of total individual and household income for small geographic areas through cross-tabulation of income with a range of census variables.

In December 2018, we'll publish Admin Based Income Statistics (ABIS)⁹. These statistics will follow on from the Income Research Outputs published in December 2016¹⁰ and October 2017¹¹. These new Admin Based Income Statistics will be defined as experimental statistics¹². As such, we'll seek feedback on the ABIS with a view to improving the statistics in terms of coverage, quality and coherence to meet user needs. We've worked closely with a range of producers of official statistics on income across the Government Statistical Service to ensure a joined-up and coherent approach.

Whilst there are still components of income missing from the 2018 ABIS income measure (in particular, self-employment and investment income), the publication will demonstrate to users how we're responding to feedback and working towards developing an income measure that meets user needs. Self-Assessment data will be incorporated into further analysis due for publication in 2019.

After the publication of ABIS in 2018, we'll seek feedback and reach out to users to establish priorities of further development work, including the potential linkage of the individual-level income data to 2021 Census records.

⁹ <https://www.ons.gov.uk/census/censustransformationprogramme/administrativedatacensusproject/administrativedatacensusresearchoutputs/populationcharacteristics/adminbasedincomestatisticsenglandandwalestaxyearending2016>

¹⁰ <https://www.ons.gov.uk/census/censustransformationprogramme/administrativedatacensusproject/administrativedatacensusresearchoutputs/populationcharacteristics/researchoutputsincomefrompayeandbenefitsfortaxyearending2014>

¹¹ <https://www.ons.gov.uk/census/censustransformationprogramme/administrativedatacensusproject/administrativedatacensusresearchoutputs/populationcharacteristics/researchoutputsincomefrompayasyouearnpayeandbenefitsfortaxyearending2016>

¹² <https://www.ons.gov.uk/methodology/methodologytopicsandstatisticalconcepts/guidetoexperimentalstatistics>

Housing

Users also gave us feedback on their need for data about property size and a more detailed breakdown of type of property. Overall, users were interested in more detailed data about properties to better understand housing stock and inform neighbourhood planning.

Local Government Association: *“Property size and type will inform the nature of housing developments by getting a better picture of current housing stock.”*

A government department: *“Deeper understanding of poverty and the economy, and the housing stock and housing need in rural areas.”*

Manchester City Council and Manchester Health and Care Commissioning: *“Size and a more detailed type of property would provide a more detailed picture of the capacity and utilisation of our existing housing stock and help plan the right type and size of housing development. The data could give us a more nuanced, localised reading of trends in our housing market if cross referenced with house price paid data for instance. Property size combined with household size might also provide some insight into areas at risk of overcrowding.”*

Cheshire West and Chester Council: *“The size and type of property would be useful for analysing overcrowding and housing conditions and again would be very useful for neighbourhood planning.”*

Property type

Users also expressed a need for data about type of property with more detail being provided than the existing census variable on type of property. This information would be used to help inform planning for future housing developments.

Pembrokeshire Coast National Park Authority: *“...It would provide useful information about our housing stock and help to identify whether we have an excess of one type of property over others. This would assist us in our collaborative work with the Local Authority around affordable housing and housing delivery and aid policy development.”*

Edmund Fallon: *“It would enable us to get a profile of the many types of property in the borough, see where they are located and help to plan for future housing developments.”*

South Norfolk District Council: *“The type of property statistics would feed into the housing market needs assessment and aiding us to develop appropriate housing policies and respond to needs of the population.”*

Central Bedfordshire Council: *“Type of property: for housing need projects, to help understand the existing stock and to make more informed plans for future developments.”*

Leicester City Council: *“Leicester faces significant housing challenges and understanding issues experienced by different property types will assist in future planning of housing services.”*

Greater Manchester Combined Authority: *“Increased detail on property type will be welcomed. Previous Census data provides breakdowns of housing types, and either number of bedrooms or number of rooms. The property type does not indicate size as it may once have done – detached houses can be smaller than terraced houses – and as such linking number and types of rooms to the property type might be beneficial.”*

Size of property

Users also identified a need for data on the size of a property. This information would be used to help understand overcrowding and quality of life.

Salford City Council: *“Size of property in square meters will be very useful. This type of data will allow the authority to establish the personal space available to residents with regard to quality of life and habitable family space, which has become a concern in recent years especially among areas of the city where large numbers of small apartments have been constructed. Existing census data allows us to identify the number of rooms / bedrooms, and although this is a good proxy it does not provide a true picture of dwelling size. For example, a large three bed flat might have more habitable floorspace than a small four bed house, although existing data is used to check against the bedroom measure of under-occupancy and overcrowding.”*

Southampton City Council: *“Size of property in square meters and type of property (more detailed breakdown) would also be useful for local government. This would give us more detail on existing properties and overcrowding.”*

These data could also be used to inform policy development and planning.

Welsh Government: *“Size of property in square meters: It would be useful to look at overcrowding/household composition, and also in relation to house prices. It might also have some practical application in funding allocations.”*

Bristol City Council: *“Size of property would be useful for analysis of space standards as part of our Local Plan and Development Management work.”*

City of London Corporation: *“To enable an analysis of the profile of different properties in terms of the size of accommodation. A simple premise we currently assume is that modern properties are of a smaller size in terms of square metres than older properties. Is this a correct assumption being made? Being able to analyse data would be useful.”*

This information was also deemed important for the emergency services to undertake risk assessments.

Shropshire Council: *“Size of property in square meters: This information has been requested in the past by colleagues from the fire service and Emergency Planning to inform their assessments of fire risk and planning for the need for smoke detectors etc.”*

We also received a response identifying how size of property would be beneficial when cross-tabulated with other census variables.

Basingstoke and Deane Borough Council: *“Size of property in square metres could be interesting if it was combined with census topics to perhaps provide average square metre figures for three-bedroom properties within a ward or perhaps square foot to population ratios. This would help to understand the variation in the housing stock and occupancy levels.”*

The data on the size of property could also be used, instead of less accurate proxy measures, in modelling land use.

A commercial sector respondent: *“We currently use data about the number of properties and the average size to calculate the size of a property in square meters when modelling land use. Having this data available would be a useful and time-saving.”*

There was also an interest in being able to compare property data across the UK to inform housing policy.

A local authority: *“A more detailed analysis of types of property, and the ability to compare this with other areas within England and Wales, would help to inform housing policy.”*

Users also stressed the need for clear metadata about the definitions used in administrative data, particularly variables derived from Valuation Office Agency (VOA) data.

Stevenage Borough Council: *“It appears that the VOA definition of property rooms is different to the previously used definitions. As Census data is a key way for us to understand property changes over the long-term, we would appreciate ONS providing clear information on comparability with previous data, particularly Census 2011.”*

Other types of data

There was also an interest in other data about housing, including property age which could be used to inform planning and service delivery.

Cornwall Council: *“Property age for example would help in targeting energy efficiency measures and communications to help maintain warm homes and reduce winter deaths, trips and falls etc.”*

Basingstoke and Deane Borough Council: *“More detail on the type of property, especially property age, could be useful in understanding the occupancy of our older dwelling stock compared to newer stock, including whether the occupants of different aged property vary significantly. This could be useful for understanding who may move into new developments.”*

Our response - Housing:

To help data users better understand housing stock, we’re exploring the potential of administrative data to produce new statistics on housing such as those that’ve been mentioned in the consultation responses. The most comprehensive housing data currently available to ONS is provided by the Valuation Office Agency (VOA). We’ll be continuing to seek access to administrative data on housing and to build knowledge about the quality of these data.

Property type

We currently plan to continue to collect information on property type in the 2021 Census, but administrative data provide an opportunity for more detailed breakdown of property type. We’ve started research into mapping property types collected in the census against information provided by VOA data and other administrative sources. We’ll also engage with users to better understand the precise detail on property type required.

Property size

We’re assessing the quality of VOA data on property size (both number of rooms and floorspace), which could support research into overcrowding and quality of life. It’s our intention to produce outputs on number of rooms in households by integrating VOA and census data.

We’ve published our research¹³ highlighting the differences in room definitions between past censuses and VOA data (for example, the VOA number of rooms excludes kitchens whilst the 2011 Census definition included them).

¹³ <https://www.ons.gov.uk/census/censustransformationprogramme/questiondevelopment/housingcommunalestablishmentsandvisitors/estimatingthenumberofroomsandbedroomsinthet2021censusanalternativeapproachusingvaluationofficeagencydata>

We've also made progress in linking VOA data with census records. Our expectation is that VOA number of rooms will meet or exceed the quality of the number of rooms information collected at past censuses. Research into the feasibility of producing outputs on property floorspace by integrating VOA and census data is still ongoing.

We're proposing that information on the number of bedrooms within households will continue to be collected through a census question. This has been shown by the 2011 Census Quality Survey to be of a high quality and will provide continuity for users interested in looking at changes in occupancy or overcrowding over time.

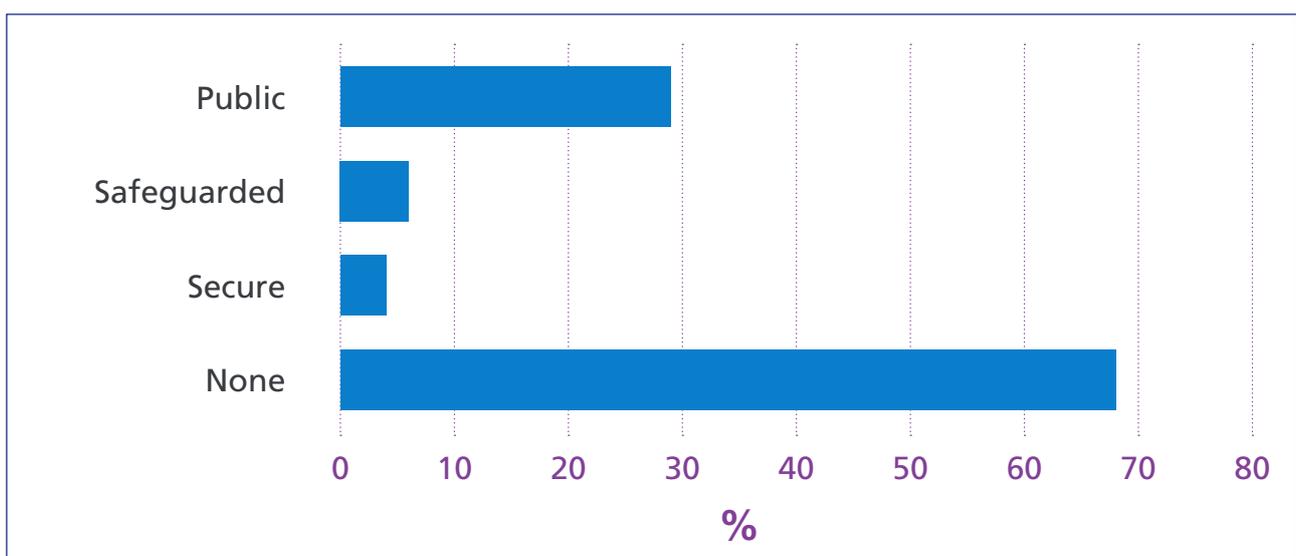
Microdata

Microdata are samples of unit-record data for whole households and individuals. They include some associated census characteristics but no information that could identify a household or individual.

We asked users which microdata samples they accessed following the 2011 Census.

Number of responses to question	184
Key findings	32% of respondents reported using microdata samples
	29% of respondents had used the public-access microdata file
	6% had used the safeguarded file
	4% had used the secure files
	Only 5% of respondents reported using more than one microdata file

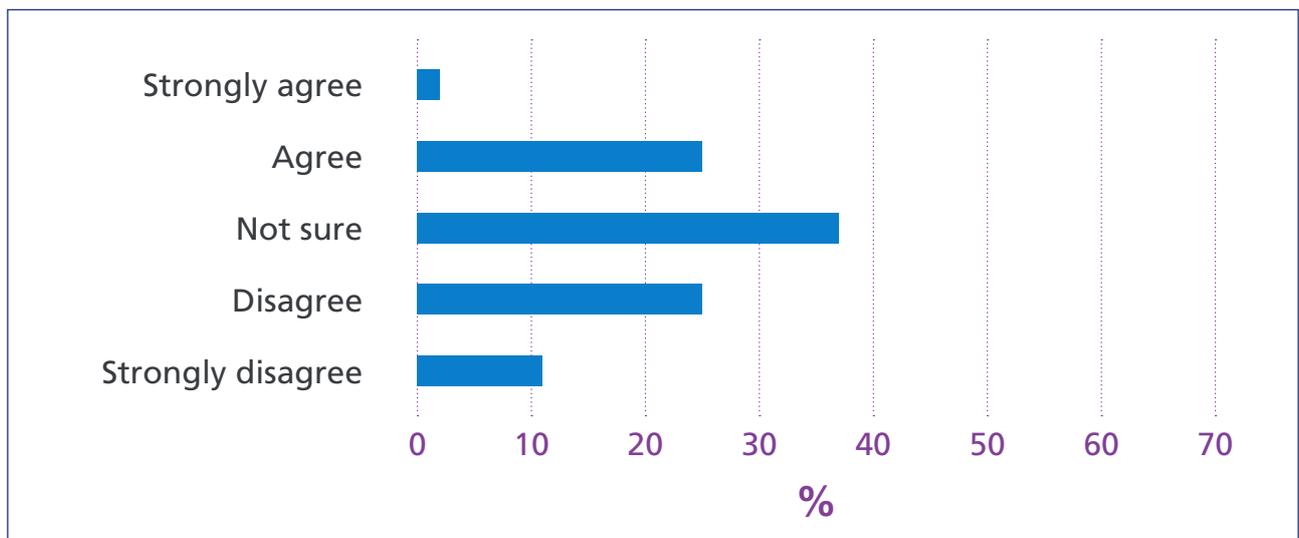
Figure 11: Responses to question “In 2011, which microdata samples did you access?”



We wanted to identify whether users felt they were sufficiently informed of the content and availability of microdata samples following the 2011 Census.

Number of responses to question	175
Key findings	37% of respondents stated they were not sure if they were sufficiently informed of the content and availability of microdata samples
	36% either strongly disagreed or disagreed with this statement
	27% strongly agreed or agreed

Figure 12: Responses to question “In 2011, you were sufficiently informed of the content and availability of microdata samples.”



There were 45 respondents who stated they were interested in contributing to the existing Microdata Working Group. We were grateful for the interest in joining this Working Group. We’ve expanded the membership of the group, with volunteers found through this consultation, to ensure we have representation from local authorities, national government, academia and the commercial sector. We took into consideration previous usage of the microdata products and the current membership of the working group, before inviting new members. We’ll ensure we continue to seek the views of other interested users and update on our progress towards designing this product.

Uses of 2011 Census microdata samples

We asked users to provide examples of research or analysis the microdata samples have enabled them to undertake. Of the 196 respondents to this consultation, 40 responded to this question.

The responses show microdata samples are being used for research and analysis across a range of sectors and themes. However, we're aware there are still a large proportion of respondents to the consultation who did not use the data.

Centre for Longitudinal Studies Information and User Support (CeLSIUS) at University College London: *"We have also used microdata samples as part of a project building a synthetic longitudinal teaching data set."*

A county council: *"Microdata is vital in order to understand patterns within communities, and small level geographies – I have used public and safeguarded access...and wanted to use secure – but never really got on with it - we used safeguarded data for our own research into origin and destination by industrial sectors, skills and occupations of those working by place of residence and vice versa."*

Greater Manchester Combined Authority: *"Microdata is being used to gain a detailed understanding of the composition of the work force practically in Regional Centre of Greater Manchester."*

A government department: *"[Microdata is being used for] rural socio-economic analysis."*

We also had several comments from users that there was a lack of awareness of the microdata samples.

A local authority: *"I was not aware of how to access the safeguarded and secure datasets."*

Staffordshire County Council: *"We were unaware of the availability of microdata."*

A local authority: *"I wasn't aware of the existence of microdata until reading the consultation document."*

Limitations of 2011 Census microdata samples

We received 33 responses from users providing evidence of limitations when using microdata samples from the 2011 Census. The limitations included accessing the samples and their design. Overall, users accepted that census microdata were valuable, however there were some aspects that users thought could be improved.

Microdata sample design

We received feedback about the need for samples that are comparable across the UK and also for comparability with 2011 census data.

The Market Research Society, and the Market Research Society's Census and Geodemographics Group and Data Analysts User Group (DUG): *"The benefits of these outputs in England and Wales would be considerably reduced if equivalent (and preferably identical, and preferably the same) outputs were not available in Scotland and Northern Ireland."*

A county council: *"I'd like to put in a plea for consistent datasets to allow comparison between census years – last time round there was an issue with the people working at or mainly at home with the 2001 census and how they were acknowledged in the journey to work stats."*

Respondents also indicated an interest in ONS producing a safeguarded household sample, which was not done in 2011.

Greater London Authority (also incorporates input from SASPAC): *"No safeguarded household sample."*

The Market Research Society and the Market Research Society's Census and Geodemographics Group and Data Analysts User Group (DUG): *"No household file was produced in 2011 which could be accessed easily (safeguarded file). Only a secure file was ever produced which is very restrictive by its nature. A household file is the most useful file for market researchers in terms of what it offers so this severely restricted the usefulness of these data."*

It was also noted that clearer metadata are required to accompany the microdata samples.

Dr Oliver Duke-Williams, University College London: *"The sample design is good; although documentation could be improved to make it cleared which variables are included (and to what level) in safeguarded verses secure files."*

Access to microdata

Several users noted there was a limited number of locations for physically accessing safeguarded and secure microdata samples. This prevented the use and ease of access of this product.

Shropshire Council: *"The location of the secure microdata would prevent a researcher from the Council being able to access the data."*

Professor Tony Champion, Newcastle University: *"The time and cost of travelling from Newcastle upon Tyne to work at VML Pimlico [ONS London]."*

Centre for Longitudinal Studies Information and User Support (CeLSIUS) at University College London: *“Secure access is a problem for both microdata and longitudinal data, specifically that there are few SRS access points. We continue to take the view that access to secure data needs to be restricted to Accredited Researchers, but would support moves to increase the number of physical locations at which access could be arranged.”*

An individual respondent: *“Access is difficult for secure if you do not live near an ONS location.”*

The process to access safeguarded data by non-government organisations and academics was also considered a hindrance to access:

RAC Foundation: *“Unreasonably onerous application procedure to access safeguarded data - no privacy risk in the data we were applying to see.”*

Our response - Microdata:

We'll continue to provide access to microdata samples for users following the 2021 Census. We've already created a microdata working group to work closely with users of microdata products to develop a specification that meets user needs. We've expanded the membership of this group following the consultation.

We want to further explore with users their priorities for the design of the samples – whether this is UK harmonisation across the 2021 Census samples or comparability with 2011 Census samples. We're also going to consider the content of the sample and whether adding more variables is feasible.

The utility should not be affected by the statistical disclosure control applied to these samples. The design of the public microdata sample will not enable individuals to be identified. For the safeguarded and secure samples, the access arrangements in place will restrict users from publicly identifying individuals included in the sample.

We'll also ensure that all our products, including microdata samples, are widely promoted to users and their benefits are explained. This includes producing detailed metadata that explain the content of the products to aid usage.

The ONS Approved Researcher process is changing¹⁴¹⁵ and will now take place through an online application portal on the GOV.UK website. It should also allow for users to connect to the Secure Research Service from a wider range of “trusted locations”. It'll also allow users to track the progress of their application online. They'll also be able to book their required training online, as well as being able to book their terminal. The Secure Research Service (SRS) will also be using new secure cloud technology that will have scalable computing power and support for data science tools.

¹⁴ <https://www.ons.gov.uk/aboutus/whatwedo/statistics/consultationsandsurveys/allconsultationsandsurveys/approvedresearcherconsultation>

¹⁵ Secure Research Service Roadshows 2018
<https://www.ons.gov.uk/aboutus/whatwedo/paidservices/virtualmicrodatalaboratoryvml>

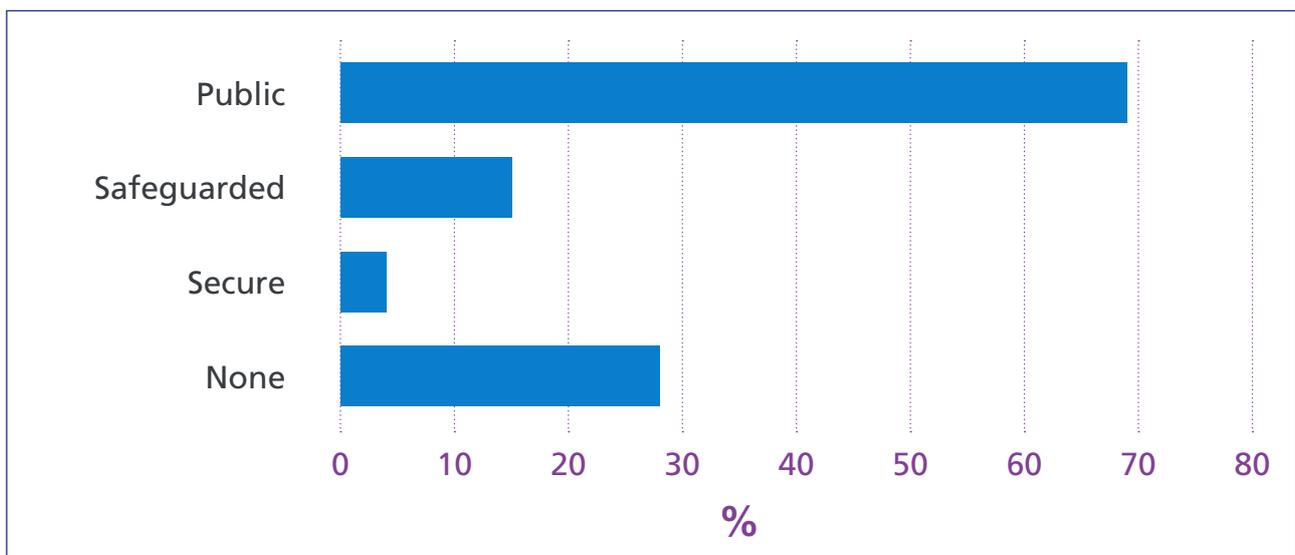
Origin-destination

Origin-destination (flow) data produced following the 2011 Census show the flows of people from one place to another.

We asked users which origin-destination data they accessed after 2011 Census.

Number of responses to question	186
Key findings	72% of respondents reported having worked with at least one origin-destination product
	69% of respondents reported having used public-access origin-destination data, 15% used the safeguarded files and 4% used the secure files
	13% of respondents reported using multiple origin-destination products
	28% of respondents stated they had not used any origin-destination data from the 2011 Census

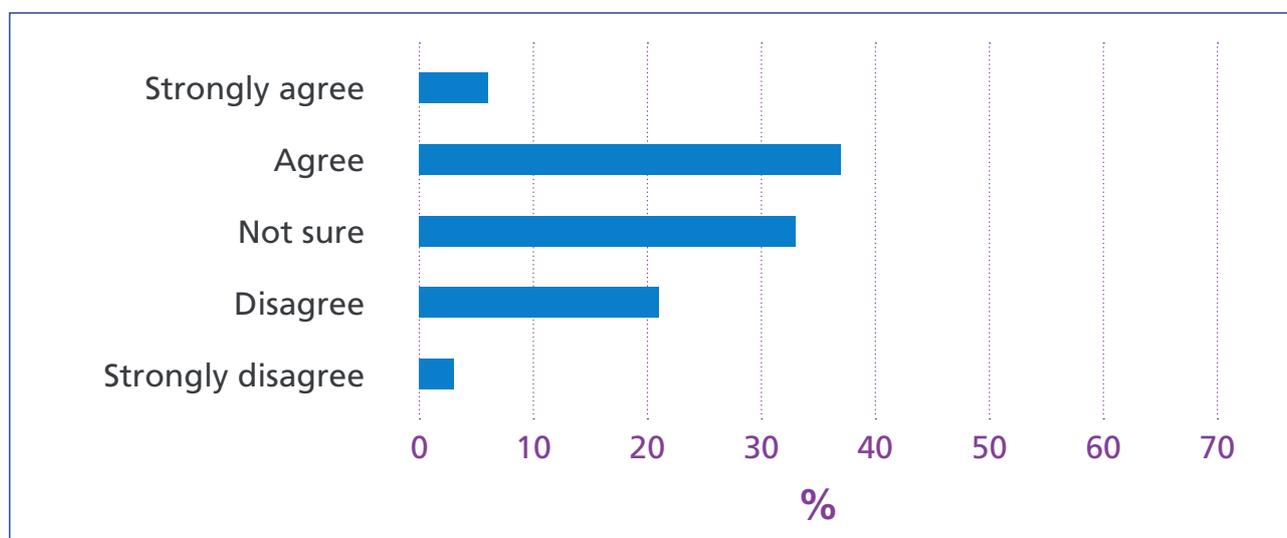
Figure 13: Responses to question: “In 2011, which origin-destination data did you access?”



We wanted to know whether users had been sufficiently informed of the content and availability of origin-destination data.

Number of responses to question	184
Key findings	43% either agreed or strongly agreed that they were sufficiently informed of the content and availability of origin-destination data
	24% either disagreed or strongly disagreed
	33% were not sure

Figure 14: Responses to question: “In 2011, you were sufficiently informed of the content and availability of origin-destination data.”



There were 53 respondents who stated they were interested in contributing to an existing Origin-Destination Working Group. We were grateful for the interest in joining this working group. We’ve expanded the membership of the group with volunteers found through this consultation, to ensure we have representation from local authorities, national government, academia and the commercial sector. We took into consideration previous usage of the origin-destination products and the current membership of the working group, before inviting new members. We’ll ensure we continue to seek the views of other interested users and update on our progress towards designing this product.

Uses of 2011 Census origin-destination data

We asked users to provide examples of research or analysis the origin-destination data have enabled them to undertake. Of the 196 respondents to this consultation, 85 responded to this question.

The responses clearly indicated that origin-destination data are highly valued by users for service delivery, informing policy and analysis by users from a range of sectors. We're pleased to hear these products are held in such high esteem and are of great value to the work of our users.

These responses show how origin-destination products have been used:

A local authority: *"Used for developing public and sustainable transport policy, particularly when connecting areas of greatest inequality to major employment sites."*

Hertfordshire County Council: *"As HCC is the local Highways authority, the MSOA level workplace flows by mode of travel (public access table WU03EW) feeds into our local transport planning process and forms part of our Town Evidence Packs...The MSOA level data is aggregated into transport plan areas to show the commuting trends across and into/out of Hertfordshire. This census OD data is vital to show these trends and although now getting out of date, it is the only comprehensive source of such data."*

City of London Corporation: *"...more focussed analysis has been undertaken for a range of specific transport related studies. The key is understanding if travel numbers and modes are increasing e.g. numbers increasing for travel of which cycling has had a large increase. It will be interesting to analyse cycling numbers from the 2021 outputs to assess whether the cycling initiatives in Greater London are delivering a change in the mode of transport."*

Cheshire East Council: *"Research into commuting flows, which shed light on (a) the scale and nature of economic, labour and housing market linkages between our local authority and its neighbouring areas and (b) the contribution that inward and outwards commuters can (or could) make to our Borough's economy."*

Bristol City Council: *"We have used the WPZ [workplace zone] travel data to inform the development of policy including a City Centre Framework, which includes development and transport proposals for the city centre. BCC have made use of the Propensity to Cycle Tool, which is based on census commuter flow data to plan future cycle infrastructure."*

Hull City Council: *"Origin destination data is a critical piece of information for us. As a really tightly drawn administrative area it really helps us to understand our relationship with the East Riding that surrounds us, understanding complex commuting patterns and our role as the sub regions urban centre."*

North Yorkshire County Council: *“Cross-border travel to work flows at local authority level and usual residence of second address owners which have been used to inform highways, economic and regional planning policy making decisions.”*

Commercial: *“Assessment of predicted travel patterns for new developments.”*

Rural England Community Interest Company: *“Travel to work flows to specific locations or types of area.”*

British Sociological Association: *“Travel to work analysis for work life studies and labour market / employment studies. Especially relevant in studies of women and work and precarious workers.”*

Limitations of 2011 Census origin-destination data

We asked respondents to describe any limitations they faced when using 2011 Census origin-destination samples, including accessing the data and the design of the products. We received 55 responses to this question. The responses indicated there were issues with the accessibility and timeliness of the release of origin-destination products in 2011.

Timeliness of release of origin-destination data

In 2011, origin-destination data were released between three and four years after census day. We had several comments from users about the reduced utility of these products due to the timing of their release.

Justine Rego: *“The main limitation was the timeliness of the release of the data; it’s about the last thing to be released and the thing that probably changes the quickest in an urban environment. So, we only used the public data as it was so out of date that it was not worth our while trying to obtain safeguarded or secure data.”*

Central Bedfordshire Council: *“This data was published a long time after 2011 (2014?), so was already losing relevancy at the point it was first available.”*

The Market Research Society and the Market Research Society’s Census and Geodemographics Group: *“The safeguarded data we accessed at the VML (single counts of workplace flows between small areas) was later made public. It would have saved us the extra time and expense of using the VML if this decision could have been made sooner. The origin-destination data was not released until 4 years after the Census in 2011. This was a significant delay and damaged using these data in a timely manner. We wanted to use these data at OA level and also found that these would not be public as was the case in 2001, but safeguarded and secure which made access much harder and more expensive. Classifying these data secure severely restricts access to these data for our users. Equivalent (and preferably identical, and preferably the same) outputs were not available in Scotland and Northern Ireland.”*

Blackburn with Darwen Borough Council: *“Origin destination outputs were only available several years after the actual census.”*

Accessibility of origin-destination products

In 2011, origin-destination data were available through public, safeguarded and secure settings. The feedback we received identified that this had caused difficulties for some users in accessing the data. This was caused by the physical location of the Secure Research Service (SRS) to access the secure data, and the restrictions placed on users wishing to access safeguarded data and export any tables created.

Bristol County Council: *“Restricted access - the publicly available data did not provide all the data that we need. Although a couple of people in BCC signed up to the UK Data Service in order to access the safeguarded SWS, this did prove to be a fundamental barrier to accessing the data. Transport planners did not embrace the data due to the complications of accessing what they needed. We are highly unlikely to ever use secure data.”*

Derbyshire County Council: *“The restricted data held on UK Data service is not as easy to access as data from NOMIS.”*

Staffordshire County Council: *“SAFE researcher training required to access data, can be quite hard to access through UKDS – ONS are much more helpful when you approach them but having to travel to either London or Portsmouth limits accessibility for most.”*

Hertfordshire County Council: *“The fact that the Safeguarded data (available via the UK Data Service) was in a different place to the publicly available data did not help. The fact that the access to the Safeguarded data via the UK Data Service was on a user by user basis added the difficulty that we could not share the data with colleagues without them registering with the UK Data Service. More guidance on this process would have been useful.”*

London Borough of Camden: *“Safeguarded access was not fit for purpose. Hosted on a different site (UK Data Archive), their data agreement meant that we wouldn’t be able to use the data freely. Not easy to get to the data in the first place, and then it was up to the user to apply disclosure control rules to the data. Spent many hours trying to get something useful out, including getting help from GLA (even they gave up) and in the end the only data that could be extracted gave nothing useful. A waste of time and energy – and a waste of the census. The data should have been made available in a format that it could have been used in; or at least a prescribed method provided.”*

Bournemouth Borough Council: *“The safeguarded data was difficult to access as the interface wasn’t very intuitive or quick to use.”*

RAC Foundation: *“Very onerous application procedure to get OD by car-driver matrices at LSOA level, even though there is no disclosure risk for this geography-mode combination.”*

Statistical disclosure control

Users also commented on the effect of the statistical disclosure control methodology on the utility of the outputs.

Greater London Authority (also incorporates input from SASPAC): *“Having to apply your own SDC was intimidating to users. Symmetric design of tables limits their use. In particular the travel to work tables included merged LAs for both origin and destination. Result was Westminster the local authority with the largest working population in England and Wales was combined with City of London, fourth largest.”*

Bristol City Council: *“By far the most important matrix is commuter flows by mode of travel. If SDC was improved by collapsing some modes into broader groups this is something that we would still find very useful, as long as the aggregations still met the needs of our transport modelling/policy objectives.”*

UK comparability

We received some comments from users about the challenges faced when trying to use cross-border flow data.

A county council: *“Issues with travel across our LA border with Scotland.”*

Dr Oliver Duke-Williams, University College London: *“The table designs were more complex than in previous years, and have not always been easy to use. The main problems I have experienced have related to the way in which tables have had disjoint geographies for flows between different member countries in the UK. These data are rather unusual in census outputs through being produced (necessarily) at a UK level, and yet pan-UK analysis is very hard in many cases.”*

Geography and variables

We also received feedback from users about the geographic levels data were provided at, and requests for more variables to be provided with the flow data.

Gwynedd Council: *“In order to improve understanding of migration flows we would ideally want this data available at OA level or below.”*

Savills: *“We were unable to get origin-destination data at the same geographic level as available from the 2001 census.”*

A commercial sector respondent: *“Level of geography not as detailed as ideal.”*

Greater Manchester Combined Authority: *“GMCA Research is interested in gross and net flows of commuters at local authority and lower levels of geography. In particular, it would be interesting to examine in more detail the opportunity to understand ‘Travel to Study’ flows data. Greater Manchester has one of the largest student populations in Europe, along with some of the largest Universities and colleges in the UK. Recent changes and potential future changes in the educational sector will have an impact on travel patterns, and so understanding a baseline for travel to study flows would help support future analysis.”*

Birmingham City Council: *“A breakdown of location of residence and place of work by occupational group and sector (SOC and SIC) would be useful.”*

Functionality of tools for accessing origin-destination data and metadata

Respondents to the consultation also identified how functionality of the 2011 Census origin-destination table design was both easy and difficult to use.

Salford City Council: *“The table design in NOMIS is very good, easy to download, and easy to use.”*

Bournemouth Borough Council: *“The safeguarded data was difficult to access as the interface wasn’t very intuitive or quick to use.”*

Greater Manchester Combined Authority: *“On several occasions the limit on cells that can be downloaded in NOMIS meant we had to split the task into several smaller jobs and then reassemble the data into a single file.”*

Commercial: *“Table design – very large to download and process, especially OA level dataset. Processing the whole UK data below local authority level is also complicated.”*

Bristol County Council: *“...an asymmetrical matrix (OA to WPZ) proved to be very complicated in practice to use for analysis, even with our extensive GIS expertise! In hindsight, a symmetrical matrix of moves between OAs would have been significantly easier to use. This is particularly the case with the creation of bespoke aggregate geographies. We need to be able to create bespoke geographies (i.e. aggregations of OAs) to match population centres across the West of England area and analyse the commuting flows between them. With an asymmetrical matrix, this basically means there is a need to create two separate lookups whilst ensuring that the two different geographies are coterminous.”*

Users also provided feedback on the quality and coherency of metadata accompanying the data to explain where which records had been included in output categories.

Huntingdonshire District Council: *“There’s been some confusion in where those working from home or who don’t have a fixed place of work are counted so that could be clearer.”*

Bristol County Council: *“Complicated output categories – the inclusion and exclusion of quasi-workplaces was not very intuitive. Even though I was the supposed ‘expert’, I had to remind myself what this actually meant every time I used the data. Excluding working workplaces - at home, offshore, not fixed, outside of UK – from the main downloads of data via NOMIS added a complication to an already complex dataset.”*

East Sussex County Council: *“Clear and consistent definitions are very important to ensure comparing like with like.”*

[Table] WU03EW – includes 16+ in employment

[Table] QS701EW – includes 16-74 year olds

In 2001, people who recorded their place of work as working mainly at or from home were considered to have their mode of travel to work as working mainly at or from home. In 2011, people working mainly at or from home could record, for example, that they travelled to work as a driver in a car or van, despite being based at home. This extra information is useful for transport planning and has been published in census table QS701EW.

There is, however, a problem in making comparisons with the 2001 Census. Therefore, the ONS published a number of additional tables, which have been generated by deriving home workers using the responses to the workplace address question (as reflected in 2001).“

Our response - Origin-destination:

We’re pleased to hear from users about the value of this product. We’re aiming to release origin-destination data earlier than we did following the 2011 Census to ensure maximum usage and relevance of the data.

Historically, it’s been difficult to find a suitable statistical disclosure control methodology that protects the confidentiality of the data, but allows data to be available to meet user needs.

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

For the 2021 Census, we've been exploring the option of protecting data by applying the cell-key perturbation method and publishing the data that pass the statistical disclosure control checks. As a result, the basic flows of individuals travelling from A to B will be considered non-disclosive, even at low levels of geography, and so could be made publicly available. When an additional variable(s) are considered, the size of the characteristic will be taken into consideration.

For small flows between areas, we'd make the size of the flow available, but not the breakdown of the characteristic variable. This method leads to a more "patchwork" approach for origin-destination data, with more breakdowns being available for larger flows without being limited by the smaller flows. This methodology was presented at the Privacy in Statistical Databases 2018 Conference in Valencia¹⁶. We hope this methodology will make more data publicly available to users.

We intend to publish publicly available origin-destination data through the ONS website and to further explore the design of the user interface allowing access to this data. We also want to understand further user needs for the functionality required to access this data easily, such as selecting the appropriate geography and variables, and download requirements.

We've already created a working group to ask for feedback on our proposed statistical disclosure control methodology and the content of this product. We've extended the membership of this group based on responses to this consultation. Through this group and other engagement, we'll develop this product, including further exploring the geographies and variables users require in the origin-destination data, whilst being mindful of statistical disclosure control parameters.

We also intend to provide clearer metadata and guidance around using the origin-destination products. As noted in the microdata section earlier in this report, we're changing the ONS Approved Research process to improve access to our sensitive data.

We're already working with our colleagues in Scotland and Northern Ireland to ensure the harmonisation of census outputs, including origin-destination products. We need to identify how we can improve on some elements of UK harmonisation and publishing cross-border flows that're more relevant for users. We also want to explore which geographies and variables are most relevant for users for cross-border flows.

ONS has also been exploring the potential of using data generated by mobile phones to estimate commuting flows and the mode of transport used for commuting, which could be produced alongside census-based outputs¹⁷. We anticipate these data could enable more timely outputs relating to commuting patterns.

¹⁶ <https://unescoprivacychair.urv.cat/psd2018/index.php?m=soon>

¹⁷ <https://www.ons.gov.uk/census/censustransformationprogramme/administrativedatacensusproject/administrativedatacensusresearchoutputs/populationcharacteristics/researchoutputsusingmobilephonedatatoestimatecommutingflows>

Geography

We asked users about two geography products – parishes and grids – that were produced for the 2011 Census, to understand the user need for these data in the 2021 Census and help inform our future research.

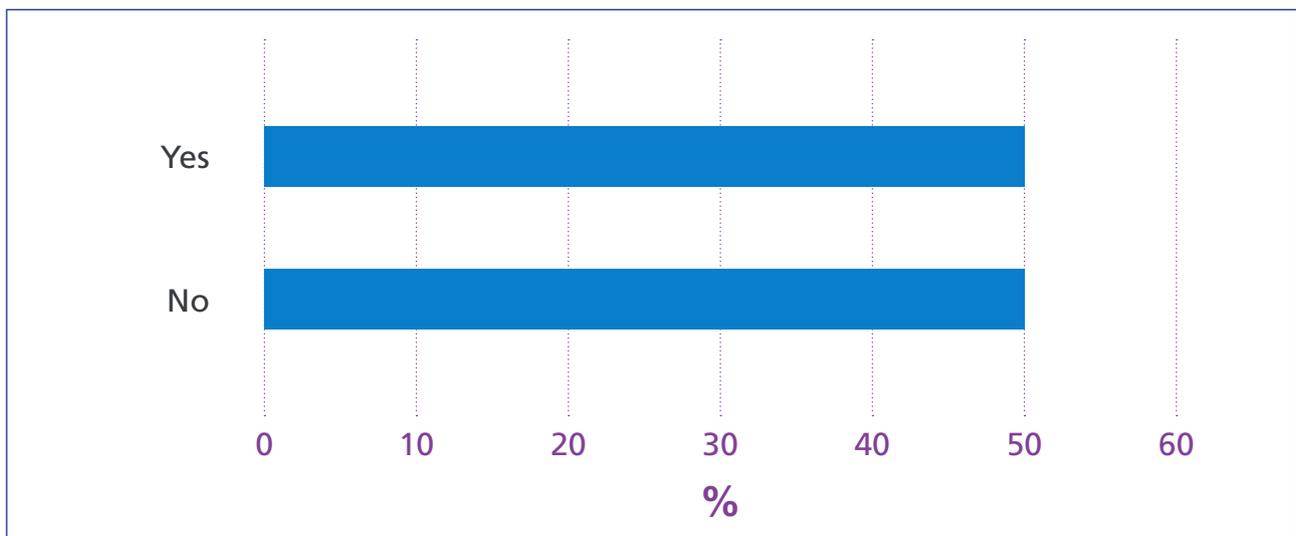
Parishes

In 2011, statistics were unavailable to 11% of all civil parishes in England, as the area covered did not contain an Output Area population-weighted centroid. We wanted to understand the importance of parish estimates for users.

We wanted to know if users had a requirement for population estimates on civil parishes.

Number of responses to question	190
Key findings	50% stated that they did have a requirement for parish data
	50% said they did not have a requirement

Figure 15: Responses to question “I have a requirement for parish data.”



We wanted to know the user requirements for parish-level statistics. It's clear that any methodology for publishing these statistics is a balance between having data available for all parishes and the number of variables that can be released.

Number of responses to question	109
Key findings	The preferred option was for users to be provided with estimates for all parishes, but some parishes will have more variables provided than others

Figure 16: Responses to question “I have a requirement for parish data.”

	Rank
Having estimates for all parishes, but some parishes will have more variables provided than others	1
Having data available for all statistical parishes even if it means that only a few variables are available	2
Having a number of variables available for parishes even if it means that not all parishes will have statistics produced	3

We asked users to provide any other comments they had on parish-level statistics.

Users noted how there's been an increased importance placed on parish-level statistics since the 2011 Census. Town and parish councils have recently taken on more responsibility for creating Neighbourhood Plans, which require parish-level data.

Cornwall Council: *“Census is the only consistent data set for Parishes and Town Councils and paints an important picture of the changing characteristics in these areas. Requirements for Parish data has increased since 2011. The introduction of neighbourhood planning through the Localism Act which came into force in April 2012 means many parish or town councils have chosen to take the lead on neighbourhood plans. These Plans can also direct service provision, for example utilities and recreational facilities which, in rural areas, may only serve a few households. The ability to profile populations at Parish/ Town Council level is therefore pivotal to this on-going role.”*

Cheshire West and Chester Council: *"I wish to emphasise the importance of Output Area and Parish level data for Neighbourhood Planning and also for analysis of very small geographical areas when looking at how to improve outcomes for local residents for examples in areas with pockets of deprivation."*

Shropshire Council: *"The recent Neighbourhood Plan legislation implemented by DCLG means many town and parish councils are considering producing Neighbourhood Plans which entail the development of a needs assessment and have statutory significance in the planning process."*

We also received two comments reiterating that data should be made available for all parishes.

A local authority: *"...the key point is that there should never be a parish or ward for which no data is published. The best outcome would be, for every parish, to publish all the data that is available for that parish. This might mean that some parishes had more data available than others, but as long as the reason why was clearly explained, that would be acceptable."*

Shropshire Council: *"The release of as much parish level data as possible would be very welcome as in a rural county like Shropshire... [We would like] an explanation of why data was not available [for parishes] and which parish it was merged with. For the 2011 Census, a word document was created for internal reference purposes... It would be useful if ONS could consult local authorities on how best to group the parishes, so that they are consistent with current town and parish groupings where possible."*

We're also aware that postcode-level data produced following the 2011 Census have been used to best-fit into other geographies such as ecclesiastical parishes.

Church of England: *"We use postcode-level headcount data and postcode centroid locations to calculate statistics for ecclesiastical parishes. We map postcodes via their centroids onto ecclesiastical parish boundaries and aggregate populations to ecclesiastical parish level. We find that this works well, and has been well accepted by our users."*

Our response - Parishes:

We recognise that the importance of parish-level statistics has increased since the last census, especially due to the requirement for data to inform Neighbourhood Plans. Users indicated this was the seventh most popular geography used (see Figure 2). We'll continue to publish population statistics on parishes following the 2021 Census.

We've started working with Ministry of Housing, Communities and Local Government, which is responsible for maintaining a list of parishes, to explore the possibility of creating a new statistical geography consisting of parish groups. This would be derived from the administrative geography of civil parishes. It would be created as an administrative geography, rather than specifically for census purposes. We can then explore whether we'll be able to release population estimates for all grouped parish councils using best-fitting from Output Areas. If this is not possible, we intend to continue to explore alternative methodologies. We'll continue to engage with users to understand the acceptability of the grouped parish council approach.

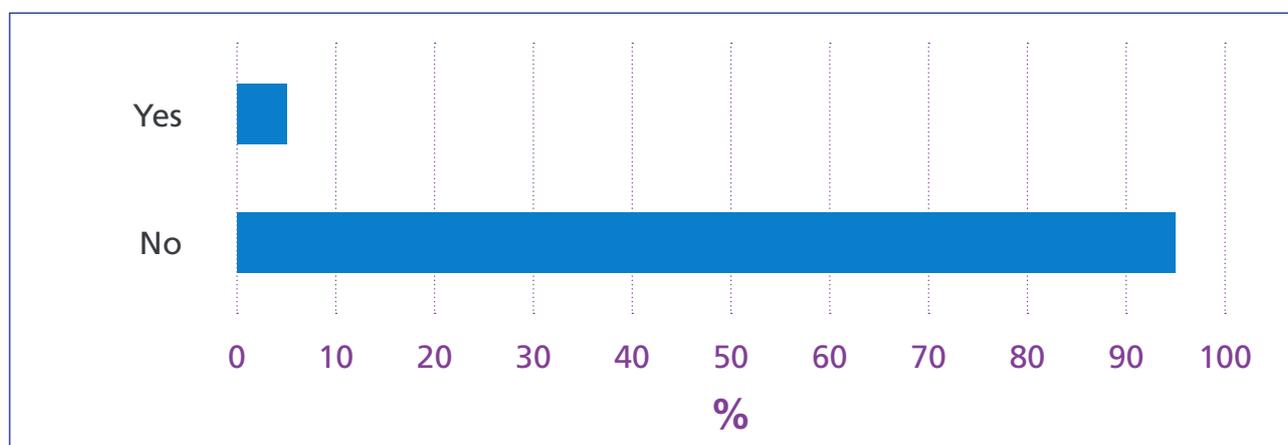
Grids

As part of our 2011 Census outputs, we released statistics on a one-kilometre grid to support analysis at the European level. We wanted to understand user requirements for statistics produced on a grid and how they might be used to support spatial analysis.

We wanted to know how many respondents to the consultation had used 2011 Census grid-square data.

Number of responses to question	187
Key findings	95% stated that they did not use grid square data from the 2011 Census
	5% stated they did

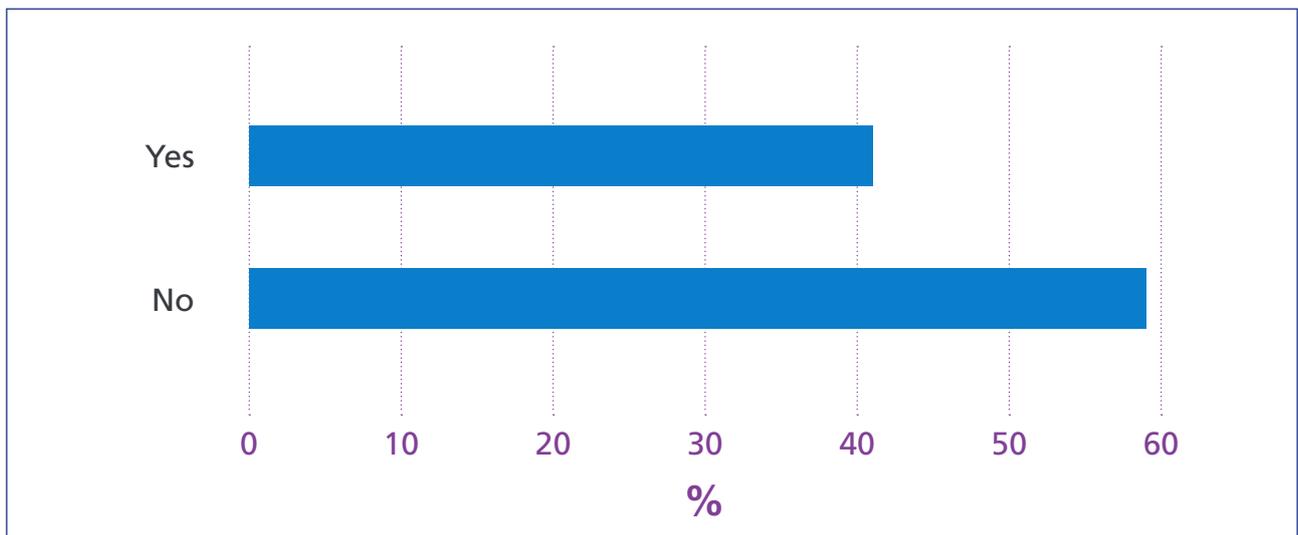
Figure 17: Responses to question "Did you use grid square data in 2011 for your work?"



We also wanted to know whether estimates produced on grids following the 2021 Census would be useful to users.

Number of responses to question	164
Key findings	59% of respondents stated that having census data on grids would not be useful to them
	41% of respondents stated that having census data on grids would be useful to them

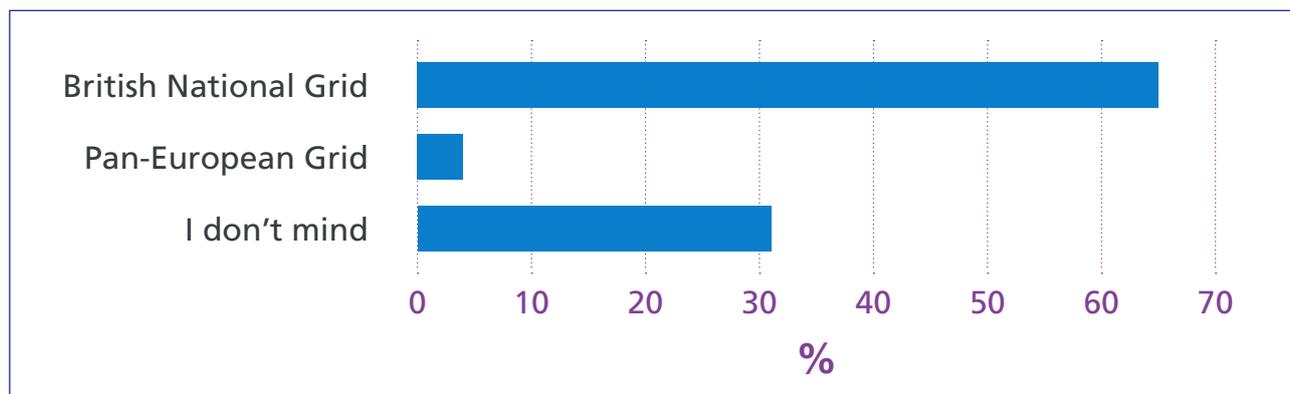
Figure 18: Responses to question “I have a requirement for parish data.”



We also asked users to indicate the grid they'd prefer to have data released on for 2021.

Number of responses to question	94
Key findings	65% of those who responded to this question stated they'd prefer to receive data based on the British National Grid
	4% said they'd prefer data released on the pan-European grid
	31% stated they did not mind which grid the data was based on

Figure 19: Responses to question “If you’ll be using grid data in 2021, which grid you would prefer to have the data released on?”



We also asked users to tell us the impact census data on grids would have on their work. We had 55 responses to this question and 35 responses to the “any other comments” questions, which included comments about grids.

Benefits of grid-based statistics

Users provided several examples where grid-based statistics have been, or would be, beneficial to their work.

Users from national government, local government and academia noted the benefit of using grids containing population data when used alongside environmental data in their work.

Department for Environment, Food and Rural Affairs (DEFRA): *“We could overlay this on our own agricultural data at 1km grid level to integrate statistics cross domains.”*

Public Health Wales Observatory: *“We would investigate how we could use grid data to produce spatial modelling of public health related environmental indicators, such as air quality.”*

Staffordshire County Council: *“We believe they will be useful for monitoring long term change and the same cell sizes allow for easy comparison. It would also allow for analysis of scientific and environment data that are not dependent on administrative boundaries, which we would be useful for teams here such as Flood Risk and Rural County.”*

Bristol City Council: *“We are likely to use the grid square data to map demographic and household data against other sources of data not mapped by OA e.g. air quality statistics, transport statistics. Statistics could be presented for example as number of cases per 1,000 households or per 1,000 people. This will help with our understanding of these areas of work and allow comparison of areas of the same geographical size.”*

Dr David Owen, University of Warwick: *“It would enable census data to be linked with environmental and satellite data.”*

Users identified how estimates produced on grids can negate the effects of irregular administrative geographic boundaries and how they change over time.

Leicestershire County Council, Business Intelligence Service: *“It would allow us to produce heat maps to help understand our communities without the distorting effect of geography. Currently we have created our own bespoke cartograms to help circumvent this issue.”*

Blackburn with Darwen Borough Council: *“Grid data provides a useful geography to present data spatially / on maps – overcoming issues around size differences for other geographies such as output areas.”*

Gwynedd Council: *“Potentially (if km grid kept constant over future years), would help in comparing trends in areas over long term periods, since this would overcome problems with administrative and OA boundary changes.”*

Greater London Authority (also incorporates input from SASPAC): *“Gridded outputs could potentially provide a consistent geography over time.”*

Users also noted how grids can provide a comparison of different areas across the UK and Europe.

Greater Manchester Combined Authority: *“The grid data could be useful for economic benchmarking of major European cities particularly comparing pre-and post-BREXIT.”*

The Market Research Society and the Market Research Society's Census and Geodemographics Group and Data Analysts User Group (DUG): *"It has the potential to provide pan-European standardised and consistent small area geography which is independent of each country's administrative geography allowing spatial analysis in an international and cross border context."*

GEOLYTIX: *"It would be super useful. We create multi-country census data packs and frequently use the Eurostat grids."*

In 2011, we only produced estimates on grids best-fitted using postcode-level data, which provided data by total population, occupied household and sex. We also received one comment about the variables that would be useful to be provided with on grids.

GEOLYTIX: *"At the moment you only publish super high-level data... It would be great if you could publish the full Census KS variables at grid. I'm sure you could come up with something to avoid disclosure either directly or by differencing. The Baltic states already do this."*

Benefits of grid-based statistics not fully realised

Despite the positive feedback on the need for grids, we also received several comments from users that they were unaware of how they'd use grid-based statistics in their work, or the benefits of these statistics.

An individual respondent: *"I didn't use the grid square data as I didn't know that it was available. May use it next time round if I know where it is and that it will be compatible with our mapping package."*

Joseph Rowntree Foundation: *"As this geographical breakdown is new to us, it would be really helpful to produce really clear user guides - with not only instructions and explanations but also some examples of how we could use these breakdowns in our work."*

Southampton City Council: *"Since we tend to use OA and LSOA data to conform to our boundaries we are unlikely to use grid data."*

Bournemouth Borough Council: *"This could potentially be useful when we map results but generally we are interested in neighbourhoods or geographies such as wards particularly when they match to natural boundaries."*

Size of grid square

We also had several comments from users about the size of grid squares, noting that a small-sized grid would be required for more densely population, often urban areas.

Bristol City Council: *“1km grid squares will enrich data in some areas where OAs are large but mean less detail for many areas. One advantage of grid squares for us will be the inclusion of blank cells with no data...At national level a 1km grid is likely to be more useful than at a local level where they might prove to be large.”*

Blackburn with Darwen Borough Council: *“Would be useful to make the grid smaller than 1km. e.g. Statistics Norway use a 250m grid across the whole country, and smaller for build-up areas.”*

London Borough of Hackney: *“1km is too large for London geographies. Please consider reducing in size for urban areas.”*

Basingstoke and Deane Borough Council: *“The usefulness would depend on the size of the grid...Grids produced at a larger scale, for example 100m could be useful for showing areas that are very low density in terms of overall population or particular characteristics.”*

Our response - Grids:

We appreciate the interest users have shown in having statistics produced on a grid. However, we're aware that there are some users who are not aware of the benefits of grid-based statistics. We plan to produce estimates on one-kilometre grid squares as we did in 2011. However, in addition, we want to demonstrate to users how estimates produced on this geography could be useful to research and analysis.

We recognise there's a strong demand for grids to be produced on a British National Grid projection. We're working with the devolved administrations (National Records of Scotland and Northern Ireland Statistics and Research Agency) to consider the issues regarding the different grid systems across the UK and to ensure UK harmonisation where possible.

We also want to further explore with users the variables that'd be most beneficial on grid squares.

Feedback we received has highlighted the benefits and limitations of a grid with only one size of grid square. We'll look to explore whether we can provide nested grid squares for areas that meet certain population thresholds, and explore the data that could be released on these differently sized grids.

ONS Geography will be holding a public consultation in 2019 to further explore user needs for geography products and services following the 2021 Census.

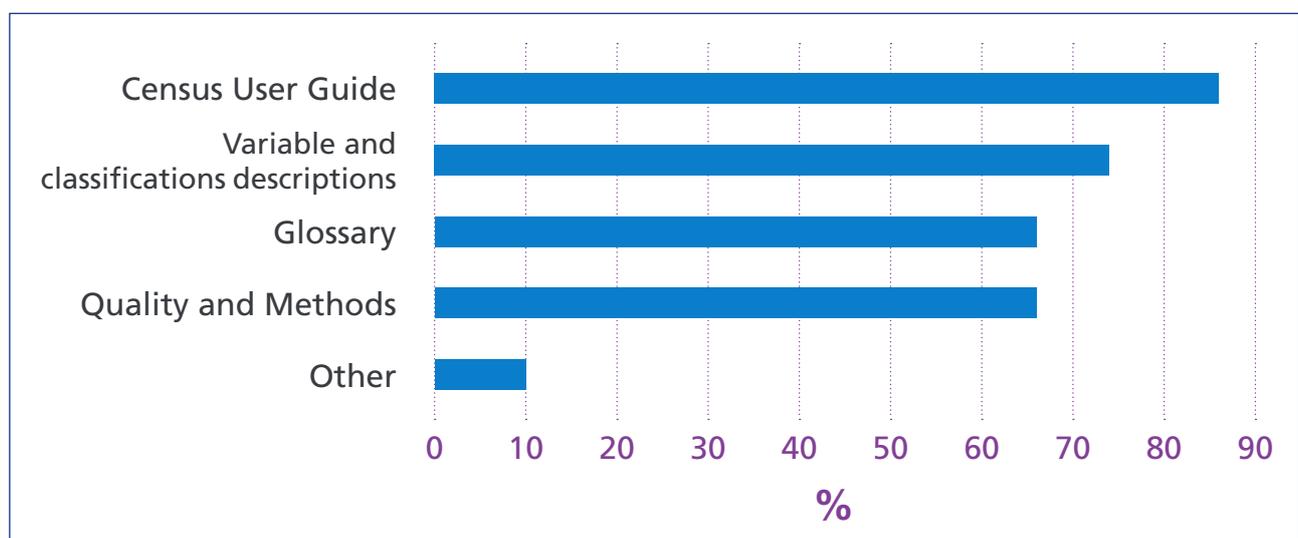
Interpreting 2021 Census estimates: supporting information

In 2011, we provided a range of documents about the information contained within the data tables, and supporting information about the quality of outputs and definitions used. This information gave context to the data, for example, how they were collected and the coverage of the data, and their publication date, descriptions and search keywords. We wanted to make sure the information we provided was useful and easy for users to understand.

We wanted to know which supporting information was used following the 2011 Census.

Number of responses to question	176
Key findings	86% of respondents had used the 2011 Census User Guide
	74% of respondents had used variable and classification descriptions
	66% had used the glossary, and quality and methods documents

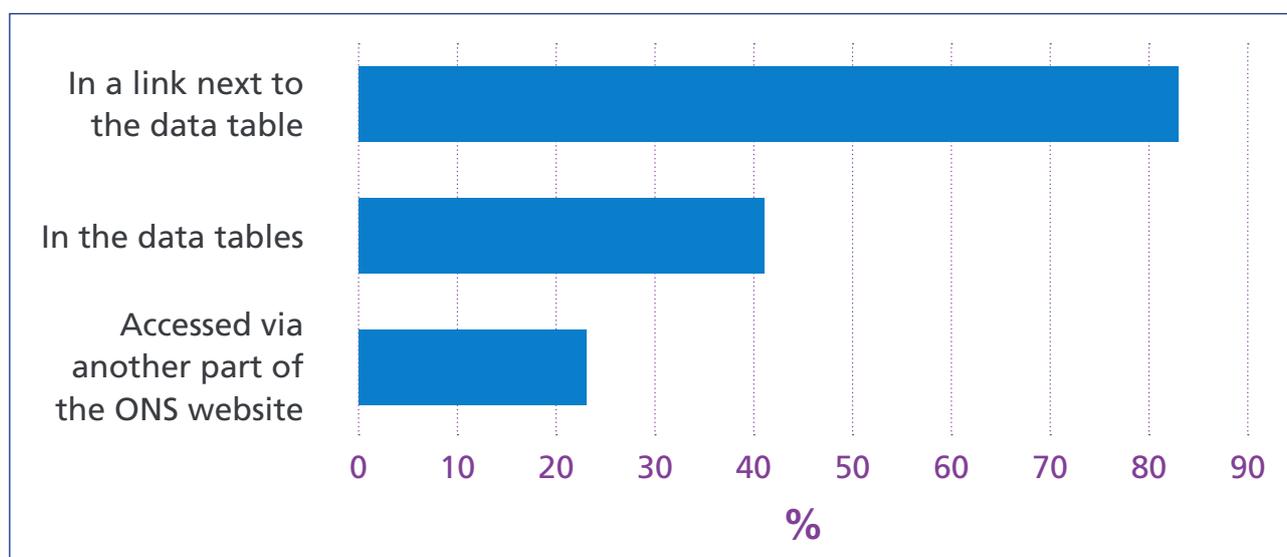
Figure 20: Responses to question “Which supporting information did you use in 2011?”



We were interested to know where users would expect to access supporting information.

Number of responses to question	186
Key findings	83% would expect to access supporting information in a link next to the table
	41% would expect to view supporting information within the data tables
	23% would expect to access supporting information from another part of the ONS website

Figure 21: Responses to question “Where would you expect to access supporting information?”



We also invited users to provide any other comments about the provision of supporting information alongside the data.

Usefulness of supporting information

Users told us they found the supporting information provided in 2011 useful.

Hampshire County Council: *“This is key. We used this enormously following 2011 and still refer to bits at times, along with the questionnaire itself.”*

Southampton City Council: *“Supporting information is a key component of the census outputs. We often refer to the supporting information and use it to support other users of the census data.”*

A local authority: *“This is vital to ensuring the correct interpretation of the datasets.”*

However, whilst users told us supporting information was important, there were only a few references to how the metadata had been used.

Blackburn and Darwen Borough Council: *“Links are easier to forward to people who have requested definitions or clarification on variable composition etc.”*

Centre for Longitudinal Studies Information and User Support (CeLSIUS) (UCL): *“Would also expect to use supporting material at other websites including academic support groups.”*

Location of metadata

We wanted to know where users expected to access supporting information, for example in the data tables, in a link next to the table, or from another part of the ONS website. We received several comments from users about where they would expect to find the supporting information alongside 2021 Census data, to ensure it was accessible for them. However, the responses did not give a strong, clear indication about where users would expect to find the information.

There were some comments that users would like metadata included in the data table they download.

City of London Corporation: *“As long as the information is available. My one comment is that internet pages change so best if it is integrated within the data tables upon data extraction.”*

Church of England: *“The guidance notes are generally very useful. Including the notes in the data tables will be useful when sharing data and outputs with others.”*

London Borough of Camden: *“The metadata is REALLY IMPORTANT so that people know what they are getting. It is useful to have the relevant metadata attached to each table, esp. variables and table populations, but also having links to the glossary so that complete understanding of the data may be appreciated.”*

Salford City Council: *“We would expect to see supporting information (such as definitions/ methodology) being included in the first tab in a downloaded file (to ensure it stays attached to the associated data when downloaded). We would not want to trawl the website for definitions and associated information that could have been provided with the actual data table, as this could result in a poor use of time resources.”*

However, we also received some feedback that users would not want metadata in the data tables.

A local authority: *"Tables will often need to be cut and pasted, so links and supporting information should not be within the tables themselves, but close at hand."*

Dudley Metropolitan Borough Council: *"It would be ideal to have the supporting information available as separate documents (as in 2011) as well as links to the relative supporting information as a link next to data tables."*

Rural England Community Interest Company: *"Could be summary info in the tables with more detailed info via a link."*

Some respondents would expect to see metadata in both the data table and accessed through another part of the ONS website.

Greater London Authority (also incorporates input from SASPAC): *"Basic metadata - population, variables used, footnotes should always be attached to the data table. But more detailed information e.g. glossary, data definitions can be separate. These could be accessed by links next to the table, or completely separate."*

A county council: *"I would expect glossary and descriptions to be alongside the data they are referring to in the online tables. The User Guide and Quality & Methods information would be separate so not to distract users not interested in this information (though the User Guide would be prominent when data is initially released)."*

There were some comments regarding the location of metadata on the ONS website, suggesting that the metadata need to be easy to find.

Lewes District Council and Eastbourne Borough Council: *"Having the links next to the data tables provides a useful and easily accessible reference point. If the supporting information was placed in another part of the website, it would be cumbersome to have to search for definitions for example, when what is required is a quick and easy way to ensure my understanding of the variables is accurate."*

Central Bedfordshire Council: *"Accessing supporting information via another part of the ONS website is not a suitable solution – the ONS website is so big it can be hard to find relevant information. Supporting information needs to be easily accessible from the same location as the table itself."*

However, we did receive comments from users who did not want metadata included in any downloadable files. This is because it would not be compatible with other functionality, such as using an API or joining with geographic files.

A local authority: *“The tables I accessed and attempted to join our GIS data to were a struggle to work with due to many tables having to be manually edited first before joining them. All description text should sit outside of the downloadable data tables making them easier to work with...blank cells are never great.”*

An individual respondent: *“It is nice having sources and dates embedded in data csv files but makes it difficult to reliably machine read or import the data.”*

East Riding of Yorkshire Council: *“Please can it be ensured that the metadata is good in APIs (as cannot influence it when used in outputs/websites).”*

Types of metadata

Users also provided us types of metadata they would expect to see provided to accompany the data.

Basingstoke and Deane Borough Council: *“Look up tables giving the names/codes of consistent tables over previous census years would be useful.”*

Birmingham City Council: *“Links to 2011 census tables should also be available for comparison.”*

Welsh Government: *“A Welsh glossary as well as an English glossary would be useful to accompany Welsh outputs so that Welsh users are using consistent terminology.”*

Greater London Authority (also incorporates input from SASPAC): *“My top requirement is for detailed descriptions of exactly how variables have been constructed and how they differ from previous censuses; analysis of this would be useful.”*

Users also expressed a requirement for metadata to be provided in plain English.

Malcolm Brown: *“The needs of different categories of users should be assessed. Some will merely want simple short explanations, others very detailed technical descriptions.”*

Darlington Borough Council: *“Accessible language for infrequent users of ONS data outputs.”*

Our response - Supporting information:

The responses clearly indicated the importance of providing metadata that are easily accessible to all users. However, they also demonstrated that there's no consensus view on how we provide supporting information to users.

We'll ensure that all supporting information is published alongside outputs and is sufficiently promoted. We aim to carry out further user research and consider the needs of different user types to inform our planning and design of metadata. We'll consider the different ways in which metadata can be provided in terms of location, quantity and detail, including how we'll provide metadata through the flexible dissemination system.

We aim to recognise the needs of different users by further assessing the impact of providing metadata alongside data tables, whilst considering machine readability when using an API and Geographical Information System (GIS) software.

We'll explore the possibility of providing information that describes the comparability of 2011 Census tables with 2021 Census data and across the UK. We'll use plain English in our metadata and follow the ONS style guide. We'll also commit to providing a Welsh translation of metadata, where required.

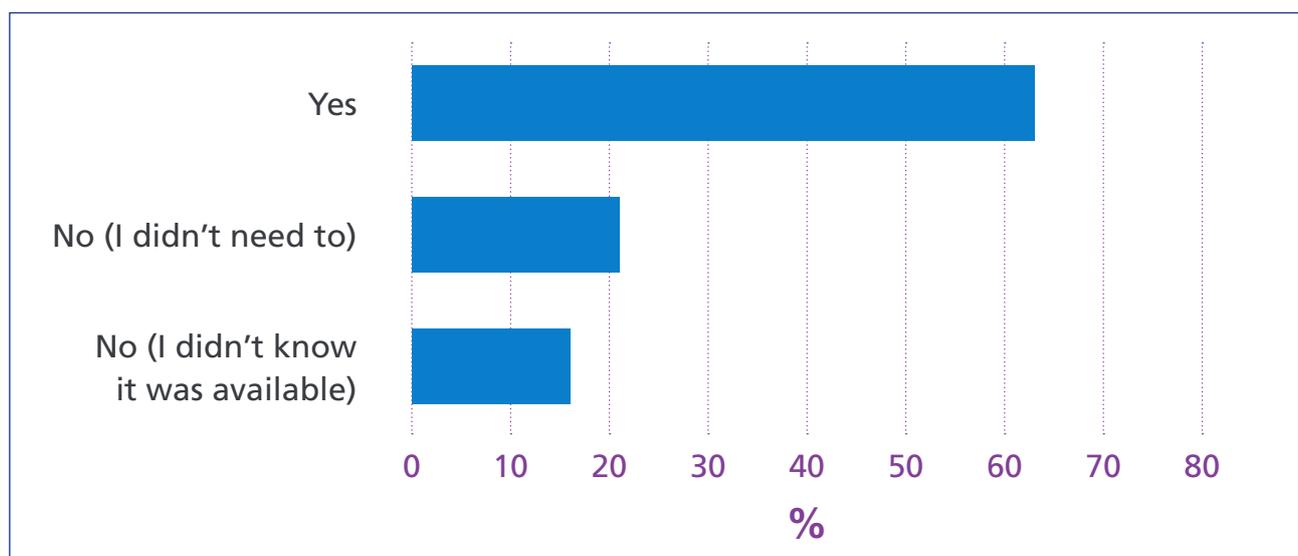
Interpreting 2021 Census estimates: wider analytical uses of census data

We're currently working to develop our analytical programme, which will support the release of 2021 Census data. We were interested to identify the user need for analysis and commentary to accompany the release of data.

We wanted to identify whether the analysis and commentary produced alongside the 2011 Census data had been utilised.

Number of responses to question	190
Key findings	63% of respondents had read the analysis and commentary published alongside 2011 Census data
	37% stated they did not read this information
	21% stated they did not need to
	16% were unaware that it was available

Figure 22: Responses to question "In 2011, did you read the accompanying analysis and commentary?"



We also wanted to know what format of accompanying material respondents found most useful.

Number of responses to question	173
Key findings	Most users stated they found accompanying commentary most useful in the format of statistical bulletins describing key outputs, followed by visualisations

Figure 23: Responses to question “What format of accompanying commentary do you find most useful?”

	Rank
Statistical bulletin describing key outputs	1
Visualisations	2
Other	3
None	4

We also asked respondents to provide any other comments about the provision of analysis and commentary. Most users preferred to read descriptions of key outputs in the statistical bulletin, as well as highlighting the importance of visualisations.

Greater Manchester Combined Authority: *“The statistical bulletin we find useful for extracting key information and summary statistics. The visualisations can save a lot of time and effort if they are a substitute for producing maps and charts ourselves. It would be useful to continue to provide summary statistics tables used for infographics / statistical bulletins as downloads.”*

Hertfordshire County Council: *“Statistical bulletins help enormously to put the results in a wider context and, because they are written for a more general audience, help us to interpret the local results and use the correct terminology when making them available to less knowledgeable colleagues. Data visualisations are extremely useful for comparing local areas with other areas.”*

However, we did receive feedback that the analysis and commentary were difficult to find for 2011 Census data in 2011 Census. It was also noted by users that the release of data should not be held up by the creation of supporting analysis and commentary.

Stevenage Borough Council: *“Found that accompanying analysis and commentary was sometimes difficult to find. Understandably it was typically focussed on interpretation at a national level, but that was often only of contextual help for us.”*

The Market Research Society and the Market Research Society’s Census and Geodemographics Group: *“We believe it is difficult for ONS to anticipate or deliver the analytical needs of a wide variety of Census data users. While some basic analysis should be delivered, the focus should be on streamlined delivery of data to allow users to do their own analysis. We would not want any delays in the release of Census outputs because the accompanying analysis is not ready in time. The speedy delivery of the outputs is of primary concern to us.”*

Our response - Wider analytical uses of census data:

We intend to provide analysis, commentary and visualisations to help support the release of 2021 Census data. We’ll build on previous analytical products and provide more detailed analysis and insight. We want to support communities to understand their needs and inform policy and service delivery. This will be done through analysing the full range of census topics alongside other relevant data sources to comment on particular themes, for example health. More information can be found in “Help Shape Our Future: The 2021 Census of Population and Housing in England and Wales”¹⁸ setting out plans for the 2021 Census of population and housing in England and Wales.

We’re working towards making our census data and analysis as engaging and accessible to as wide a range of audiences as possible. We’re also aware the 2011 Census analysis and commentary was difficult to find. We’ve improved the design and functionality of the ONS website, which users should find easier to navigate.

We’ll follow the Code of Practice for Statistics, to ensure estimates are published in a timely manner and accompanied by analysis that will provide insight and tell our users about the society we live in. We’ll outline our plans for this closer to the release of 2021 Census data.

¹⁸ <https://www.ons.gov.uk/file?uri=/census/censustransformationprogramme/legislationandpolicy/helpshapeourfuturethe2021censusofpopulationandhousinginenglandandwales/englishlanguageversion.pdf>

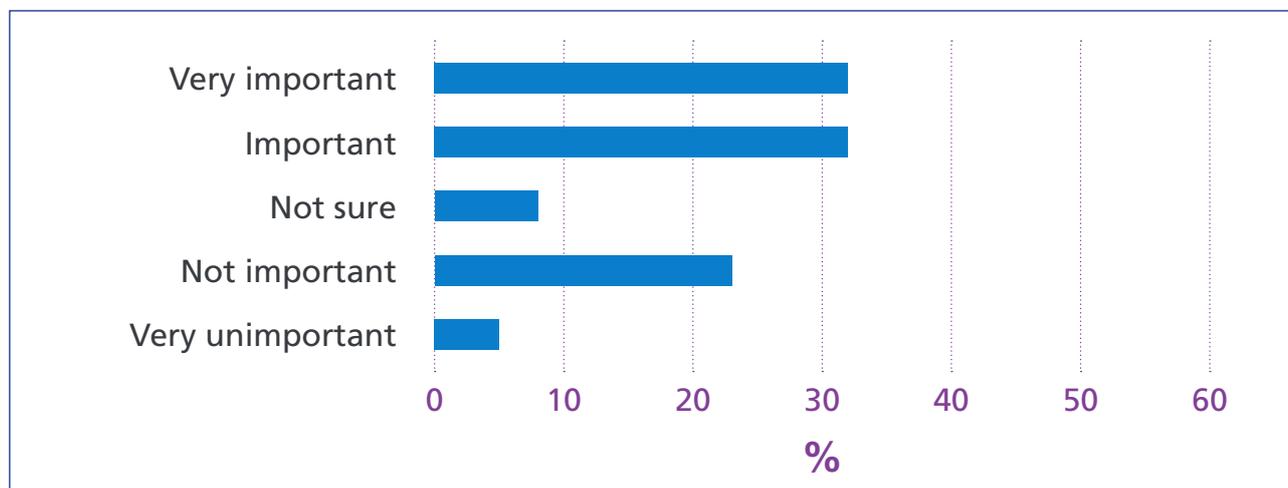
Comparability of 2021 Census estimates: UK harmonisation and UK statistics

ONS has a responsibility for disseminating 2021 Census statistics for the UK, where data collected permits this to be done. As stated in the Statement of Agreement for the conduct of the 2021 censuses across the UK, we'll aim to achieve harmonisation of 2021 Census output content, where possible.

We were interested to identify how important it is for users to be able to access data from England, Scotland, Wales and Northern Ireland.

Number of responses to question	192
Key findings	64% responded that being able to access data from each of England, Scotland, Wales and Northern Ireland was either important or very important to their work
	28% stated it was either unimportant or very unimportant to their work
	8% stated they were not sure

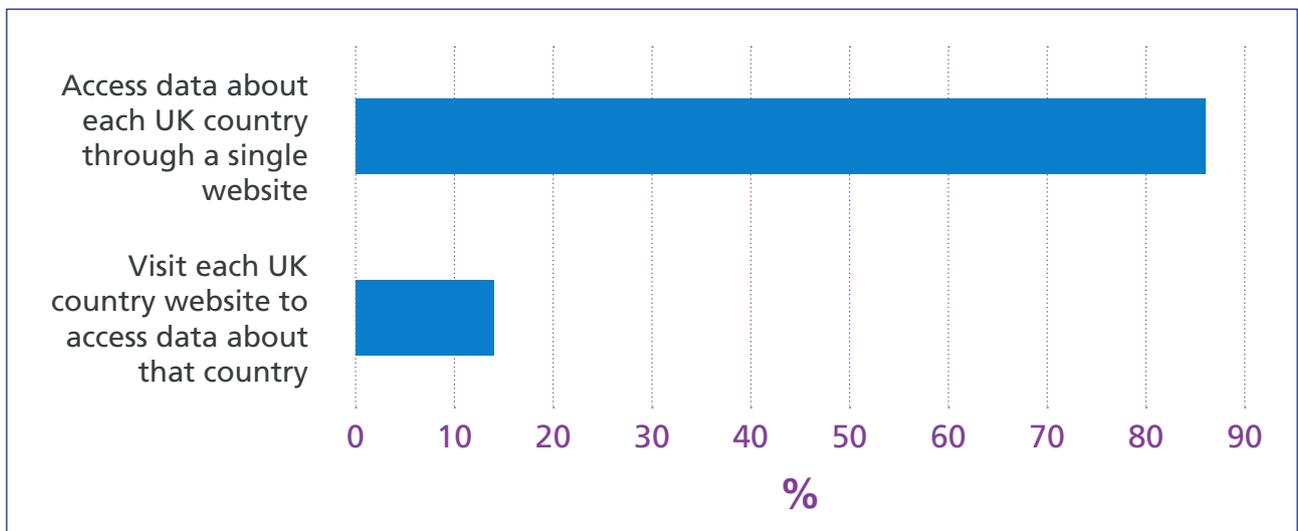
Figure 24: Responses to question “How important is being able to access data for England, Scotland, Wales and Northern Ireland for your work?”



We also wanted to identify how users would expect to access UK data.

Number of responses to question	185
Key findings	86% of respondents stated they would expect to access data about each UK country through a single website
	14% stated they would expect to visit each UK country website to access data about that country

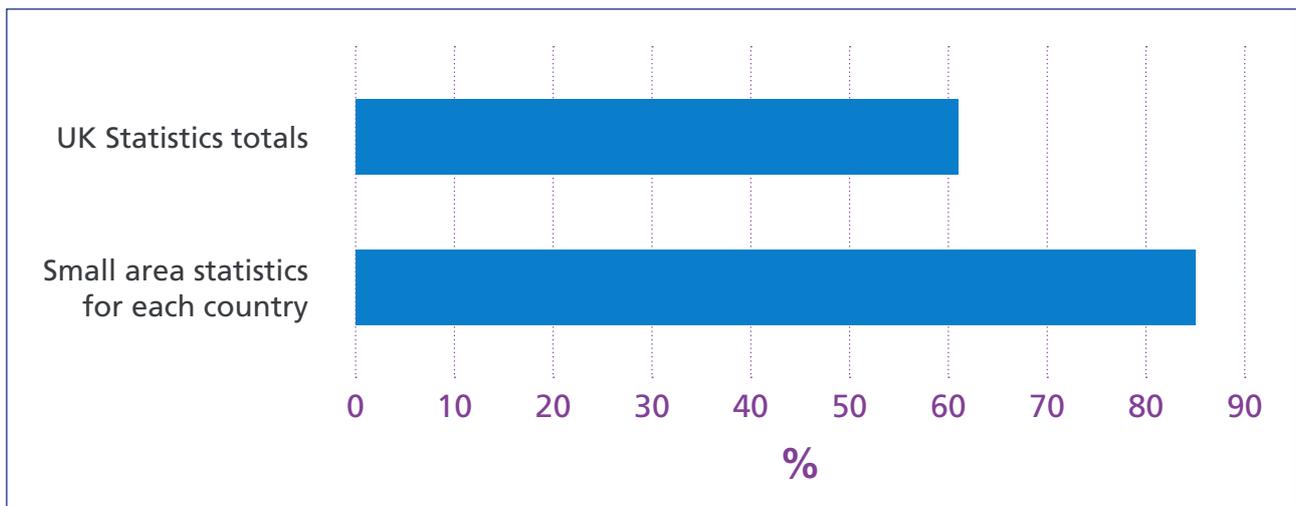
Figure 25: Responses to question “How would you expect to access UK data?”



We also wanted to know the geographic level of UK data users were interested in.

Number of responses to question	178
Key findings	85% of respondents indicated they were more interested in receiving statistics for each country, rather than the UK as a whole
	61% of respondents indicated that they would be interested in UK statistics totals
	47% of respondents expressed an interest in both forms of data

Figure 26: Responses to question “Are you interested in UK statistics totals, small area statistics for different geographies across England, Wales, Scotland and Northern Ireland, or both?”



We wanted users to tell us about the impact of UK statistics on their work and how they might expect to access the data.

Uses of 2011 Census UK data

Users demonstrated to us how they had utilised UK statistics following the 2011 Census.

UK data were often used as a benchmark for analysis to understand the local picture in the context of the UK.

Worcestershire County Council: *"We need access to statistics on countries mainly for comparison purposes and to benchmark local values. This would usually either be UK or England. For context, it is important for us to know how Worcestershire and the areas within are doing compared to the national average for a large number of census outputs."*

Oxford City Council: *"We provide statistics about Oxford that use national UK statistics as benchmarks."*

Salford City Council: *"We are frequently asked for the national comparator whenever we support colleagues and stakeholders with their work. A UK comparator often useful as well as an England comparator depending on work requirements and data availability (we often get asked for a UK comparator and if it isn't available then we endeavour to provide a close match for comparison purposes for example GB or England). A lack of comparator options often makes it difficult to provide the requester with the information they want to get a true picture of how the local situation compares with the rest of the country for benchmarking and service planning purposes. Also, the UK provides the largest population as the highest available geographic level with which to compare data and trends. Examples where we have compared Salford with the UK covers a wide range of work including the State of the City Report 2016, ward & neighbourhood profiles, Orthodox Jewish Housing Study, Salford Workforce Equality Report..."*

Rotherham Metropolitan Borough Council: *"Comparison with the UK average rather than English average was beneficial for demonstrating the need for higher education provision, as qualification levels are higher in Scotland, raising the average."*

House of Commons Library: *"The need for UK statistics reflects the fact we have a UK Parliament. The Library provides research and analysis to MPs from all parts of the UK. The availability of statistics on a UK wide basis allows us to provide essential context to our users, so that MPs from all parts of the UK can understand the position of their constituency – or any other relevant area – relative to the rest of the country. Where possible, we already provide statistics on a UK wide basis in many House of Commons Library briefing papers."*

Welsh Government: *"Comparing to other parts of the UK is useful in terms of policy context and development."*

Public Health Wales Observatory: *“Accessing UK data gives us the ability to make comparisons between population health in Wales and UK/English regions.”*

Users also needed UK data to compare their area with other similar areas across the UK.

Malcolm Brown: *“I appreciate that Census outputs will not be the same in each UK Country but I often find it more useful to, say, compare Cornwall with rural local authorities in Wales and Scotland than metropolitan areas of England. It should be made as easy as possible to do that.”*

A government department: *“Each country has their own definition of rural and conducts analysis according to their priorities. With Brexit, it will be important to have an understanding of issues affecting rural areas across the whole of the UK. Some analysis has already started to be done at the UK level, and this will become more important in a post-Brexit environment.”*

UK data are also used to inform policy development and monitoring for UK-wide issues.

Home Office: *“The devolution of powers within the UK means analysis by region of the UK is key in policy development and monitoring. However, Immigration and Visa policy is primarily developed on a UK wide basis and therefore it is vital to have comparable UK aggregate data. The delay in being able to access UK wide data means that we have not been able to utilise it to its full potential. Timely UK wide 2021 Census data would be very useful to policy development and fulfil data obligations as set out below.*

Member State obligations to EC partners (Eurostat) require submission of UK level data for publication in order to deliver comparisons across Members States within the EU. This would include cross tabulations of: passports held, housing tenure, occupation and other variables. Failure to provide this data may result in financial sanctions for the UK. UK-wide data also enables wider international comparisons to be made, as most other OECD (Organisation for Economic Co-operation and Development) countries collect information on citizenship and/or nationality in their Censuses.”

Accessing UK data

We also received feedback from users about how they would expect to access UK data. The comments indicated users would like UK data to be available from a single point of access. This reflected the evidence presented in figure 25, which indicated that 86% of respondents to that question would like to access data via a single point of access.

A county council: *“Being a Local Authority, we need to place ourselves in the national context and having access to all UK data is useful. The problems with getting Scottish and Northern Irish data makes this work difficult. Being a rural*

authority, it can often be the case that the areas of the UK that are most similar to us are Scottish and this data being a step more difficult to access has an impact on the ease of doing our work."

Bristol City Council: *"Bristol is one of the 10 Core Cities in Great Britain. As such we often make comparisons with the other cities in this group which now includes Glasgow. A few times we haven't been able to make the necessary comparisons easily. It would be simplest if all the census data was available on one web site."*

RAC Foundation: *"We are a UK charity, working with small area statistics. Multiple websites add to our labour, and different outputs mean you have to go do more work to ensure comparability."*

UK harmonisation

Users also provided valuable feedback on the harmonisation of UK outputs, particularly surrounding the definitions and outputs released.

Salford City Council: *"We fully agree with the principle that all census outputs should have full UK harmonisation unless there is a strong case not to, such as different 'school qualifications', the 'welsh language' or the local question on 'when someone arrived in Northern Ireland'."*

The Market Research Society and the Market Research Society's Census and Geodemographics Group and Data Analysts User Group (DUG): *"We create UK-wide geodemographic classifications and derived data products, e.g. CACI's Acorn classification, Experian's Mosaic classification, and Beacon Dodsworth's P2 People & Places classification. It is very important to us that we have UK wide harmonised statistics in terms of data definitions and outputs released. Of particular importance to us are consistent outputs across England, Wales and Scotland."*

The Market Research Society and the Market Research Society's Census and Geodemographics Group and Data Analysts User Group (DUG): *"For the 2011 Census we had UK harmonised univariate outputs (which was a great success), but this was not the case for multivariate outputs. For 2021, we would like UK harmonised univariate and multivariate outputs."*

In 2011, the UK univariate outputs were released around 6 months after each Census office had released their own univariate outputs. Timeliness is key to our users so we took the earlier outputs and created UK outputs ourselves before the UK release. Would not want any delays in the release of outputs to occur while waiting for one of the three Census offices to be ready to release their data for the UK output to be built.

We would like an agreed list of what is to be included as UK outputs before release so we can plan accordingly.

For 2011, ONS produced a daytime population and workplace bases, but this was not reciprocated for the rest of the UK, and the outputs produced were not that extensive particularly in the case of the daytime population. For 2021, we would like and hope that the daytime population and workplace bases are produced for the UK with an extensive set of outputs. We would also like these to be available via the flexible dissemination system.”

Our response - UK harmonisation and UK statistics:

We welcome this feedback to help establish the need for UK data and harmonised outputs, where possible. Feedback we received also illustrates the importance of UK data being available from a single point of access to improve efficiency in users’ work.

We’ve already started work with our colleagues in Scotland and Northern Ireland to enable the harmonisation of variables and timely dissemination of UK data, and our preferred option for doing so is through a single point of access for users. We’re working to ensure a technological solution is available to allow this to happen.

Feedback we received showed the importance of having UK data to provide a benchmark for analysis, and there was some user need for comparisons across the UK at smaller geographies.

We’ll be conducting further user research to identify the variables, geographies and population bases required in our UK releases and how users would expect to access UK data.

Where the data collected is consistent across the census offices of the UK, we’ll make results for Great Britain and the UK available. In recognition of the importance of providing consistent outputs across the UK, the National Statistician and the Registrars General for Scotland and Northern Ireland have signed an agreement on the conduct of the census in 2021. This includes, in particular, the production of harmonised statistics with specific attention being given to common questions, population bases, definitions, disclosure control methodologies and output formats¹⁹.

We’ll lay first results from the 2021 Census for England and Wales before Parliament, present them to the National Assembly for Wales and publish them.

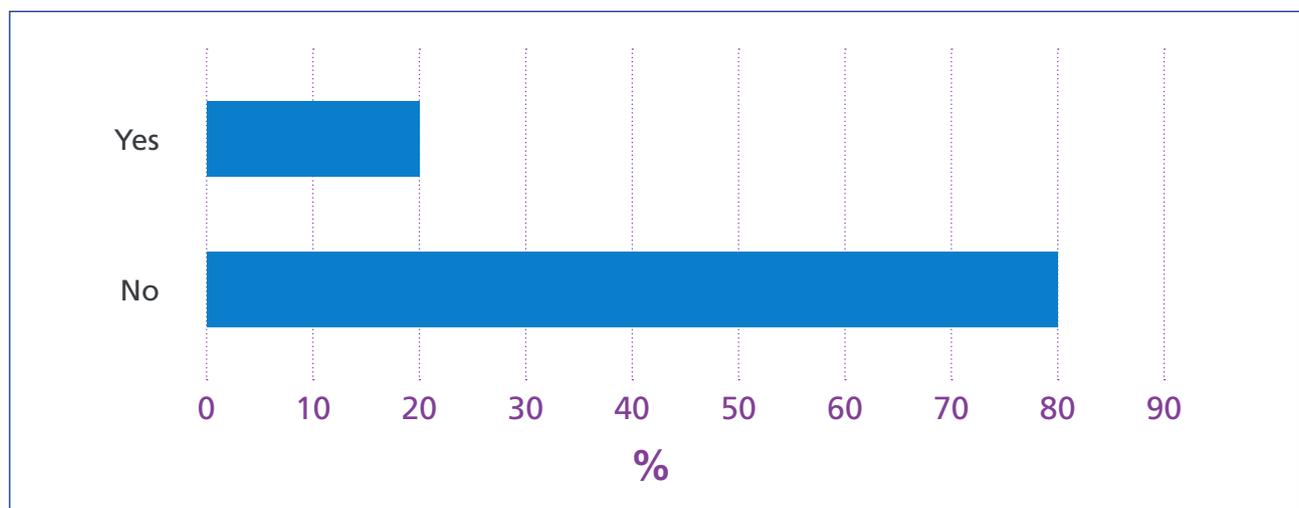
¹⁹ <https://www.ons.gov.uk/census/censustransformationprogramme/legislationandpolicy>

International comparability of 2021 Census results

We wanted to identify where users had compared UK estimates to international datasets following the 2011 Census and the importance of this for their work.

Number of responses to question	191
Key findings	80% said they had not compared UK estimates to international estimates
	20% said they had

Figure 27: Responses to question “Have you compared UK estimates to international datasets?”



We asked users to provide examples of where UK data has been compared to international datasets.

There were several responses received showing the need for international comparisons of census data, particularly in the changing European and global political and economic climate.

Southampton City Council: *“When doing a Building European Environmental and Maritime Skills (BEEMS) project we compared the UK, Southampton and England with other maritime based economies.”*

A local authority: *“We have used Eurostat and US government open data to compare the ageing population of our local authority against other global areas.”*

Shropshire Council: *“Shropshire and the LEP are very focused on the County’s role within the global economy, not just the UK and Europe. Often international comparators are required for setting the context and for monitoring Shropshire economic health. This information informs the LEP / Shropshire’s Economic Growth Strategy and is included in the Shropshire Economic Profile.”*

Home Office: *“Home Office Science: Migration and Border Analysis is the UK’s national contact point for the European migration network (EMN). The purpose of the EMN is to provide up-to-date, objective, reliable and comparable information on migration and asylum to inform policy development at National and EU levels. Census 2011 migration fields (such as passports held) has enabled the UK to fulfil its EMN obligation to produce national thematic studies to feed into European-level synthesis reports. For example, in 2014 the EMN produced a study looking into migrants’ access to social security and healthcare. Census 2011 data was used to inform the picture for the UK identifying, for example, differences in employment rates by passports held and allowing Europe wide comparisons.”*

Shropshire Council: *“It is likely that as the UK leaves the European Union over the next two years and establishes trade agreements with new global partners, there will be an increased demand for intelligence that enables comparison with international datasets.”*

Greater Manchester Combined Authority: *“Since the EU Referendum we have increased our economic monitoring work. International comparisons, particularly to European comparator cities, is useful in this line of work and will be something we will be doing increasingly in the coming months and years.”*

Our response - International comparability of 2021 Census results:

Responses we received illustrate the importance of users being able to compare UK estimates to international datasets. ONS recognises the importance of producing UK-level data that are internationally comparable.

A large part of the UK-level data produced from the 2011 Census was required by European Union (EU) census legislation. We’ll need to review our exact approach as the outcome of the EU withdrawal negotiations become clearer, but there are some clear advantages of producing the UK-level census statistics required by the EU. They help to meet demand for UK-harmonised outputs, provide comparable data over time and with other European countries, and form the basis of data provided to other international organisations such as the United Nations. Many of the data required are also produced anyway.

The innovative statistical disclosure control methodology that will underpin the flexible dissemination system has been agreed as the default method for protecting hypercubes (datasets) across the EU in the 2020 census round²⁰.

We'll continue to meet our international commitments to supply census data to the United Nations Statistical Commission following the 2021 Census.

²⁰ https://ec.europa.eu/eurostat/cros/system/files/recommendations_for_the_protection_of_hypercubes.pdf

Next steps

We'd once again like to thank everyone who gave us their views by responding to the consultation. The evidence you've provided is invaluable in helping us to identify the user need for 2021 Census outputs and prioritise our future research. The final design and content of 2021 Census outputs and overall dissemination approach will be driven by statistical and technical methodologies, but most importantly user needs.

In 2019, ONS will carry out a rehearsal of the collection, data processing and dissemination components of the census. We're going to be assessing the interfaces and dependencies, processes, methods and systems relating to outputs and dissemination to ensure we're prepared and co-ordinated to deliver the 2021 Census outputs. This will include testing the functionality of the flexible dissemination system with dynamic statistical disclosure control operating on the ONS technical platform. It will also include the creation of microdata and origin-destination products. We will not be publishing any data externally from the 2019 Rehearsal.

We recognise there's a broad spectrum of census data users, with many different interests and needs. For this reason, it will not be feasible for us to consult with all census data users to the same extent and level of detail. However, we'll continue to engage with users in a variety of ways throughout the design and dissemination of outputs.

We also plan to continue to test the functionality of our proposed dissemination mechanism, and make decisions informed by extensive user research. User research involves engaging with users, often on a one-to-one basis, to identify their specific needs and issues, before collating their comments and feeding back to the project team to inform future design.

If you'd like to give us more feedback, please contact the 2021 Census Outputs and Dissemination team at census.outputs@ons.gov.uk or visit [our webpages](#)