

Admin-based accommodation type statistics for England and Wales, feasibility research: 2011

Further research demonstrating the improved potential of Valuation Office Agency (VOA) data to provide detailed information on accommodation type, and examining how VOA data compares with the 2011 Census data, and a request for feedback on the usefulness of these statistics.

Contact:
Shannon Bull, Joseph Herson,
Emily Mason-Apps, Stephan
Tietz
admin.based.
characteristics@ons.gov.uk
+44 1329 444528

Release date:
10 March 2022

Next release:
To be announced

Notice

10 March 2022

The [previous article](#) in this series is located in a different area of the website.

Table of contents

1. [Main points](#)
2. [About our transformation research](#)
3. [Statistical quality and coverage of VOA data](#)
4. [Mapping VOA accommodation type](#)
5. [Imputing missing VOA accommodation type](#)
6. [Comparing VOA accommodation type to 2011 Census accommodation type](#)
7. [Future developments](#)
8. [Glossary](#)
9. [Data sources and quality](#)
10. [Related publications](#)

1 . Main points

- This work demonstrates the viability of Valuation Office Agency (VOA) data to provide detailed information on accommodation type, which supports resource allocation and housing strategies.
- We successfully linked 97% of 2011 Census households to VOA property data; the property type of “flats within converted/shared houses” saw the largest increase (10 percentage points) of successful linkage in comparison with our original research, ensuring that all types of property are well represented in our research.
- We made use of additional information on VOA data (for example, VOA dwelling code) and can now identify 73% of “flats above/within commercial buildings”.
- We demonstrated that VOA property type values that are missing or discontinued, and cannot be mapped onto an accommodation type, can be imputed from historical census records; this improved the coverage by 0.9%.
- We propose to introduce an accommodation type for annexes as these are clearly recorded by the VOA but not by the census; this affects 0.1% of addresses.
- The overall agreement between VOA accommodation type and census accommodation type is 89%, which is an increase of 13 percentage points compared with the original research; comparatively, the Census Quality Survey found a 92% agreement rate with 2011 Census for accommodation type.
- We identified a preference for 2011 Census respondents to define their property as “semi-detached” instead of “end-terrace” and, similarly, for census respondents to favour “purpose-built flats” where VOA identifies “flats within converted/shared house”; using VOA accommodation type could provide more consistency than self-reported census data.

Disclaimer

These research outputs and the linked [Valuation Office Agency \(VOA\) property characteristics data](#) used for the analyses are not official statistics on accommodation types in England and Wales. Rather, they are published to allow users to compare the VOA data with the 2011 Census as a potential alternative approach to collecting information on accommodation type. These outputs should not be used for policymaking or decision-making. Official statistics about property type include [Council Tax statistics](#) and [2011 Census accommodation type statistics](#). Accommodation type statistics for Census 2021 are also part of phase 1 of [our release plan proposals](#).

2 . About our transformation research

At the Office for National Statistics (ONS), we are [exploring the use of administrative data on housing](#), which might remove the need for us to collect data on accommodation type through a census or surveys. It also has the potential to provide new breakdowns of accommodation types not previously available. This may help meet user needs for more detailed categories of accommodation type, which can further support central government resource allocation and local government housing strategies ([see the Census 2021 consultation](#)).

This research is a progression of [original research](#) that compared property type data recorded by the Valuation Office Agency (VOA) with data collected by the 2011 Census in England and Wales. Our original method for deriving accommodation type using VOA data was unable to provide comparable information with the 2011 Census category for dwellings “above or within commercial buildings”. This new research addresses this gap as well as dealing with properties where administrative data on their accommodation type are not available.

This research forms part of [our population and social statistics transformation programme](#), which aims to provide the best insights on population, migration and society using a range of data sources. The findings will form part of the evidence base for the National Statistician’s recommendation in 2023 on the future of population and social statistics in England and Wales.

3 . Statistical quality and coverage of VOA data

Census households were linked with Valuation Office Agency (VOA) addresses using a unique property reference number (UPRN). We have improved the way we assign UPRNs to records on both data sources (see [Section 8](#)). The overall linkage rate for households was 97% for both England and Wales. That is an increase of one percentage point for England and three percentage points for Wales in comparison with our original research.

The proportion of census households that could be linked to VOA data increased for all accommodation types. High linkage rates (between 97% and 99%) were achieved for “detached”, “semi-detached” and “terraced (incl. end-terrace)” properties as well as “purpose-built flats”. “Flats within converted/shared house” saw the largest increase, rising by 10 percentage points to 74%. The high linkage rate across all accommodation types means that all types of properties are well represented in our research.

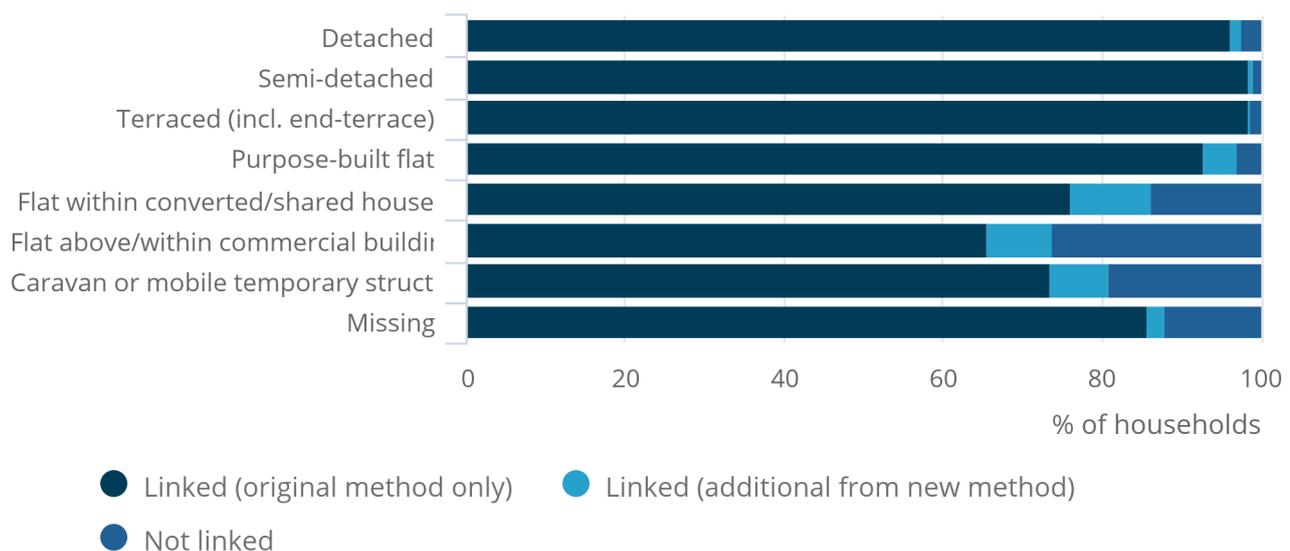
The distribution of households that linked for each census accommodation type are shown in Figure 1. The “missing” category is where the householder did not respond to the question on the census about accommodation type or gave an invalid response ([see Section 8](#)).

Figure 1: Proportion of census households that linked to VOA data by census accommodation type

Proportion of census households that linked to VOA data, by census accommodation type, England and Wales, 2011

Figure 1: Proportion of census households that linked to VOA data by census accommodation type

Proportion of census households that linked to VOA data, by census accommodation type, England and Wales, 2011



Source: Valuation Office Agency and Office for National Statistics

4 . Mapping VOA accommodation type

The 2011 Census refers to the type of property (for example, a detached house or flat) as “accommodation type”, whereas the Valuation Office Agency (VOA) data records “property type” and “dwelling code”. VOA property type is a similar concept to census accommodation type, but it has more detailed subcategories that can mostly be mapped onto census accommodation type. The VOA dwelling code provides additional information to enable the identification of “flats above/within commercial buildings”. This is an improvement to the original mapping method used in our [previous research](#). Our new mapping of VOA property type and VOA dwelling code onto VOA accommodation type can be found in [Section 8](#).

Three levels of agreement have been defined to qualify if VOA accommodation type and census accommodation type agree with each other for a given household.

Levels of agreement

Full agreement

The census accommodation type and VOA accommodation type are the same. For example, “detached” on the census and “detached house” on VOA data.

Partial agreement

There are either understandable reasons why the census response might not be the same as the VOA accommodation type, or we do not have enough information to determine if they are the same.

For example, “semi-detached house” on the 2011 Census and “end-terraced house” on VOA data, or “terraced (incl. end-terraced)” on the 2011 Census and “houses in a cluster” on VOA data.

Contradiction

The two accommodation types are in contradiction. For example, “caravan or mobile temporary structure” on the census and “purpose-built flat with lift” on VOA data.

Original mapping of VOA accommodation type

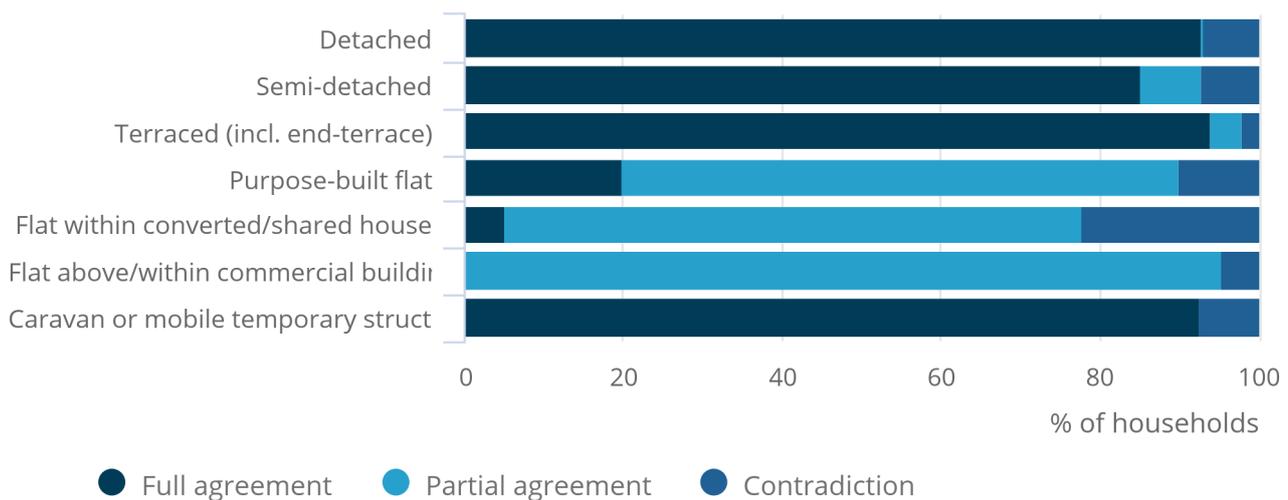
Figure 2 shows the agreement rates for the original mapping of VOA accommodation type, with an overall agreement rate of 76%. Since conducting the original research, we have worked with the VOA to better understand the definition of VOA property type categories and how they relate to census accommodation types. This has allowed us to define the levels of agreement more accurately. As previously stated, without the use of VOA dwelling code, it was not possible to identify “flats above/within commercial buildings”, and we have tweaked the levels of agreement to reflect this more clearly. Additionally, when presenting the original mapping within this publication, we have been cautious and classified “flats within converted/shared house” and “purpose-built flats” as partial agreement. This is because we cannot say for certain that these are not “within/above commercial buildings”. This explains why the agreement rates for the original mappings are lower here than in the [previous publication](#).

Figure 2: Agreement rates for the original VOA accommodation type mapping by census accommodation type

Agreement rates for the original VOA accommodation type mapping, by census accommodation type, England and Wales, 2011

Figure 2: Agreement rates for the original VOA accommodation type mapping by census accommodation type

Agreement rates for the original VOA accommodation type mapping, by census accommodation type, England and Wales, 2011



Source: Valuation Office Agency and Office for National Statistics

New mapping of VOA accommodation type

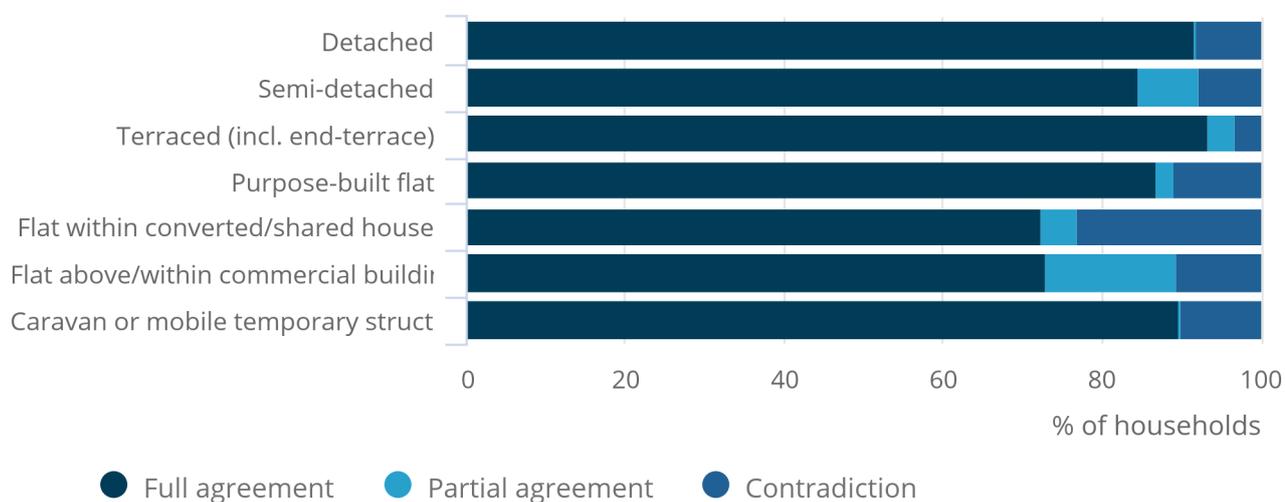
Figure 3 shows the agreement rates for the new mapping of VOA accommodation type. The new mapping method has higher levels of agreement for all accommodation types compared with the original mapping, with an overall agreement rate of 88%, which is an increase of 12 percentage points. The overall agreement rate is 88% for England and 90% for Wales. Importantly, the new mapping can accurately identify 73% of “flats above/within commercial buildings”.

Figure 3: Agreement rates for the new VOA accommodation type mapping (without imputation) by census accommodation type

A stacked bar chart showing agreement rates for the new VOA accommodation type mapping (without imputation)

Figure 3: Agreement rates for the new VOA accommodation type mapping (without imputation) by census accommodation type

A stacked bar chart showing agreement rates for the new VOA accommodation type mapping (without imputation)



Source: Valuation Office Agency and Office for National Statistics

There are still households with “partial agreement” or “contradiction”. The agreement rate can be further increased by creating a new VOA accommodation type for “annexes” as these are not recorded on the census, but they are by the VOA (see [Section 6](#)). The agreement rate can also be further increased by imputing a value for VOA accommodation type where it is either missing or is derived from a discontinued VOA property type (see [Section 5](#)).

5 . Imputing missing VOA accommodation type

Some census households cannot be assigned a Valuation Office Agency (VOA) accommodation type for two reasons. Firstly, this happens if the value for VOA property type is missing. Most missingness in VOA data relates to historical records rather than newer addresses. Secondly, no census accommodation type can be fully mapped for addresses that have a VOA property type of “houses in a cluster”, “bungalows in a cluster”, “houses of unidentified type”, “bungalows of unidentified type” or “flats of unidentified type”. These VOA property types are discontinued, which means they are not assigned by the VOA to new addresses.

Missing and discontinued VOA property types are primarily associated with older addresses in VOA data. Therefore, it is possible to permanently reduce the level of missingness in the derived VOA accommodation type by imputing (or copying) the accommodation type information available from the 2011 Census for the approximately 200,000 affected records. Table 1 shows the number of properties that linked to census households that either had a missing or discontinued VOA property type. Supplementing the data with historical census records would occur for a relatively small number of cases (0.91% of census records).

Table 1: Proportion of addresses with missing and discontinued VOA property types that linked to the 2011 Census

| Household | VOA property type missing (%) | VOA property type discontinued (%) | VOA accommodation type to be imputed (%) |
|--|-------------------------------|------------------------------------|--|
| Detached | 0.99 | 0.05 | 1.04 |
| Semi-detached | 0.27 | 0.16 | 0.42 |
| Terraced (incl. end-terrace) | 0.22 | 0.87 | 1.09 |
| Purpose-built flat | 0.77 | 0.13 | 0.90 |
| Flat within converted/ shared house | 1.69 | 0.18 | 1.87 |
| Flat above/within commercial building | 5.79 | 0.18 | 5.96 |
| Caravan or mobile temporary structure | 2.48 | 0.04 | 2.52 |
| Overall | 0.60 | 0.31 | 0.91 |

Source: Valuation Office Agency and Office for National Statistics

Figure 4 shows the agreement rates for the new mapping of VOA accommodation type using the imputation method described previously. As expected, the agreement rates have increased further across all census accommodation types. The overall agreement rate is now 89% compared with 88% without imputation and 76% for the original method. The overall agreement rate is now 89% for England, compared with 88% without imputation, and 90% for Wales, which is the same as without imputation.

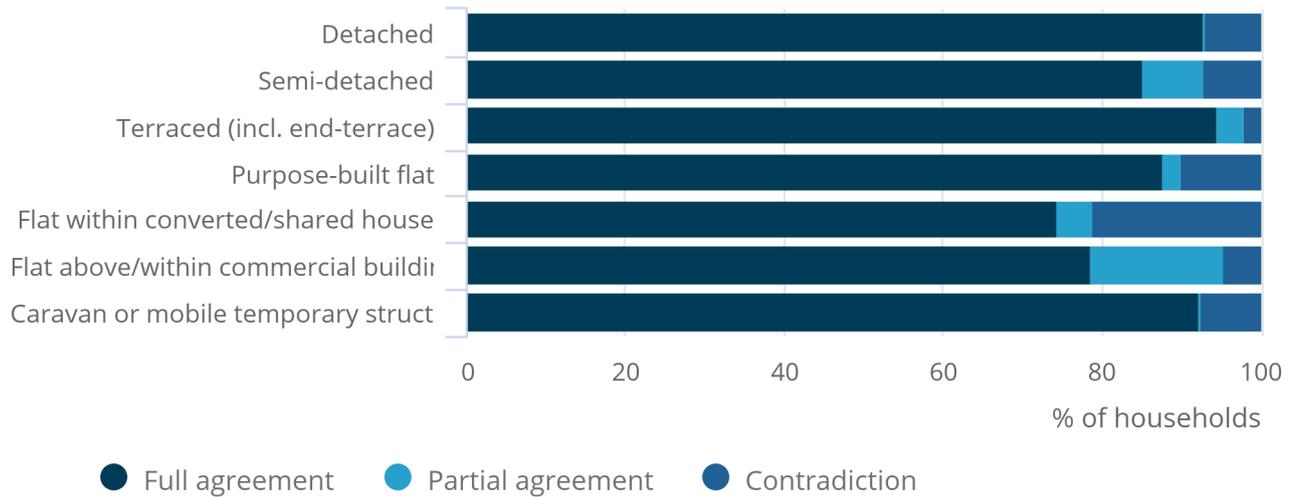
Comparatively, the quality of the census responses for accommodation type was measured by the [2011 Census Quality Survey \(CQS\)](#) at 92%. The CQS was conducted after the 2011 Census to assess the quality of census responses. See [Section 6](#) for a comparison of VOA accommodation type with 2011 Census responses.

Figure 4: Agreement rates for the new VOA accommodation type mapping (with imputation from historical census data) by census accommodation type

Agreement rates for the new VOA accommodation type mapping with imputation from historical census data, by census accommodation type, England and Wales, 2011

Figure 4: Agreement rates for the new VOA accommodation type mapping (with imputation from historical census data) by census accommodation type

Agreement rates for the new VOA accommodation type mapping with imputation from historical census data, by census accommodation type, England and Wales, 2011



Source: Valuation Office Agency and Office for National Statistics

A further benefit of the new mapping method and using historical census data for imputation is the very high level of completeness compared with collecting this information through the census. The 2011 Census had to impute accommodation type for over 516,000 households (2.29%). In contrast, our method would only require further imputation for approximately 7,000 addresses (0.03%). Future work will explore using Census 2021 accommodation type data.

6 . Comparing VOA accommodation type to 2011 Census accommodation type

Splitting out annexes for VOA accommodation type

[Section 5](#) discussed a number of Valuation Office Agency (VOA) property types that cannot be mapped to an accommodation type, and we propose to use historical census data to impute their value. The final VOA property type that cannot be mapped onto census accommodation type is “annexes”, and we propose to introduce this as a new accommodation type for reporting.

Table 2 presents the proportion for each census accommodation type in comparison with VOA accommodation type using the new mapping (see [Section 4](#)), imputation of missingness (see [Section 5](#)) and splitting out annexes. The largest difference in distributions of accommodation type would be “semi-detached” and “terraced (incl. end-terrace)”, with a difference of negative 2.98 and positive 3.53 percentage points respectively. The difference can be explained by census respondents favouring selecting “semi-detached” instead of “end-terrace”. It is worth noting that the classification as “end-terrace” from the VOA data is likely the one that a trained surveyor would choose. For all other accommodation types, the difference was smaller (between negative 0.95 and positive 0.65 percentage points).

Table 2: Distribution of census and VOA accommodation type and their difference

| Household | Census accommodation type (%) | VOA accommodation type (%) | Percentage point difference (VOA-Census) |
|---------------------------------------|-------------------------------|----------------------------|--|
| Detached | 22.88 | 21.96 | -0.92 |
| Semi-detached | 31.69 | 28.71 | -2.98 |
| Terraced (incl. end-terrace) | 25.39 | 28.92 | 3.53 |
| Purpose-built flat | 15.73 | 14.78 | -0.95 |
| Flat within converted/shared house | 3.27 | 3.92 | 0.65 |
| Flat above/within commercial building | 0.79 | 1.40 | 0.61 |
| Caravan or mobile temporary structure | 0.26 | 0.27 | 0.02 |
| Annexe | N/A | 0.05 | 0.05 |
| Total | 100.00 | 100.00 | |

Source: Valuation Office Agency and Office for National Statistics

Differences at local authority level for VOA accommodation type

Figure 5 shows the percentage point difference between VOA accommodation type and census accommodation type by local authority. A positive number indicates that there is a higher proportion of the corresponding accommodation type on the VOA data, while a negative number indicates a higher proportion for that accommodation type on the census.

Figure 5: Differences between census and VOA accommodation type by Local Authority

Download the data

[.xlsx](#)

As discussed previously, there is a preference for census respondents to define their property as “semi-detached” instead of “end-terrace”. This effect can be seen nationally but is most notable in Greater London. The largest effects of this are seen in Barking and Dagenham, where “semi-detached” is 10.3 percentage points higher for census accommodation type and “terraced (incl. end-terrace)” is 15.7 percentage points lower.

Cities across England and Wales tend to have more properties being listed as “purpose-built flat” by the census and “flats within converted/shared house” by VOA. This effect is especially noticeable in London. This could be because of the high percentage of renters in London, with renters potentially being less aware of the history of the property than a homeowner or a surveyor collecting information reported to the VOA. For example, the London Borough of Kensington and Chelsea has 15.2 percentage points more “purpose-built flats” using census accommodation type and 13.7 percentage points fewer “flats within converted/shared house”. The opposite effect is seen in the north of England.

Overall, the percentage point differences between the census and the VOA at a local authority level are low. The exceptions to this are “detached”, “semi-detached” and “terraced” properties in Greater London and “purpose-built flats” and “flats within converted/shared house” in cities across England and Wales. Using VOA accommodation type could provide more consistency than the self-reported census data.

7 . Future developments

This research demonstrates the viability of producing accommodation type statistics for residential properties in England and Wales using Valuation Office Agency (VOA) property type and dwelling code. In future, we may be able to produce more frequent, census-like statistics by accommodation type.

We will continue to explore how we can further improve admin-based accommodation type statistics through:

- establishing if we can identify the newly introduced [Census 2021 accommodation type](#) of “flat, maisonette or apartment which is part of another converted building”
- developing a way to identify occupied residential addresses from administrative data sources

Feedback

We welcome feedback on the method used to produce the admin-based accommodation type statistics and the planned future developments. We are very interested in understanding what accommodation types are likely to be of interest in the future to inform policies, target schemes and monitor changes over time. This information will help us to ensure we meet user needs where possible. Please email your feedback to admin.based.characteristics@ons.gov.uk. Please include “Housing” in the subject line of your response.

8 . Glossary

Accommodation type

Accommodation type is a measure of the property. For example, a flat or terraced house. See “Census accommodation type” and “Valuation Office Agency (VOA) accommodation type” in the glossary for more detail.

Census accommodation type

The 2011 Census has seven different accommodation types. They are:

- a whole house or bungalow that is detached
- a whole house or bungalow that is semi-detached
- a whole house or bungalow that is terraced (including end-terrace)
- a flat, maisonette or apartment that is in a purpose-built block of flats or tenement
- a flat, maisonette or apartment that is part of a converted or shared house (including bedsits)
- a flat, maisonette or apartment that is in a commercial building (for example, in an office building, hotel, or over a shop)
- a caravan or other mobile or temporary structure

Imputation

Imputation is the process of identifying and treating errors in data. For this research, errors in VOA data refer to the property type value being missing. There are different imputation methods available; we propose to impute (or copy) the missing information from historical census records. This is a type of cold-deck imputation.

Missingness or missing values

Missingness occurs when a household has no accommodation type. For this research, this happens when the VOA address linked to the census household has no VOA property type or the VOA property type cannot be mapped to a census accommodation type.

Unique property reference number (UPRN)

A unique property reference number (UPRN) is a unique identifier for every address in Great Britain and is allocated by local government and Ordnance Survey (OS).

VOA accommodation type

The VOA accommodation types are derived from VOA property type and VOA dwelling code to resemble the census accommodation type as closely as possible.

Detached

“A whole house or bungalow that is detached” in the 2011 Census dataset is captured by the following VOA property types:

- detached house
- detached bungalow

Semi-detached

“A whole house or bungalow that is semi-detached” in the 2011 Census is captured by the following VOA property types:

- semi-detached house
- semi-detached bungalow

Terraced

“A whole house or bungalow that is terraced (including end-terrace)” in the 2011 Census is captured by the following VOA property types:

- mid-terraced house
- end-terraced house
- mid-terraced bungalow
- end-terraced bungalow

Purpose-built flat

The “flat, maisonette or apartment that is in a purpose-built block of flats or tenement” category from the 2011 Census dataset is captured by the following VOA property types if the VOA dwelling code is not “flats and maisonettes above and rear of shops and similar commercial premises”:

- purpose-built flat with lift
- purpose-built flat without lift
- purpose-built maisonette with lift
- purpose-built maisonette without lift

This is a change to the original mapping method to enable the identification of “flats above/within commercial buildings” (see glossary entry for that term).

Four additional VOA property types can be mapped without the use of VOA dwelling code. These are:

- purpose-built, self-contained flat in two-storey, detached block
- purpose-built, self-contained flat in two-storey, semi-detached block
- purpose-built, self-contained flat in two-storey, mid-terrace block
- purpose-built, self-contained flat in two-storey, end-terrace block

Flat within converted/shared house

The “flat, maisonette or apartment that is part of a converted or shared house (including bedsits)” category from the 2011 Census dataset is captured by the following VOA property types:

- flat with lift in converted former commercial or industrial buildings
- flat without lift in converted former commercial or industrial buildings
- maisonette with lift in converted former commercial or industrial buildings
- maisonette without lift in converted former commercial or industrial buildings
- non-purpose-built flat, not self-contained
- non-purpose-built maisonette, not self-contained

VOA property types of “non-purpose-built flat, self-contained” and “non-purpose-built maisonette, self-contained” are also included if the VOA dwelling code is not “flats and maisonettes above and rear of shops and similar commercial premises”. This is a change to the original mapping method to enable the identification of “flats above /within commercial buildings” (see glossary entry for that term).

Flat above/within commercial building

The “flat, maisonette or apartment that is in a commercial building (for example, in an office building, hotel, or over a shop)” category from the 2011 Census dataset is captured by the following VOA property types if the VOA dwelling code is “Flats and maisonettes above and rear of shops and similar commercial premises”:

- purpose-built flat with lift
- purpose-built flat without lift
- purpose-built maisonette with lift
- purpose-built maisonette without lift
- non-purpose-built, self-contained flat
- non-purpose-built, self-contained maisonette

We were not able to identify this category as part of our original mapping method, which did not use VOA dwelling code.

Caravan or other mobile temporary structure

The “caravan or other mobile temporary structure” category from the 2011 Census dataset is captured by the VOA property type of “caravan” and “mobile home or house boat”.

Annexe

“Annexe” is not a category from the 2011 Census dataset, but it is a new category we propose for the VOA property type of “annexe”.

VOA dwelling code

The VOA has 51 different dwelling codes. For the mapping of VOA accommodation type only, the value “flats and maisonettes above and rear of shops and similar commercial premises” was used.

VOA property type

The VOA has 29 different property types plus a further three categories for “unidentified” houses, flats and bungalows. They are referenced in the glossary entry for “VOA accommodation type”, and you can also find a list in our [original publication](#).

9 . Data sources and quality

2011 Census data

The Office for National Statistics (ONS) currently conducts a census every 10 years. Census statistics help paint a picture of the nation and how we live. They provide a detailed snapshot of the population and its characteristics, underpinning funding allocation to provide public services.

This research uses the census accommodation type for each household collected by the 2011 Census. The census defines a household as “one person living alone, or a group of people (not necessarily related) living at the same address who share cooking facilities and share a living room or sitting room or dining area”.

Most residential addresses in England and Wales are used by a single household, but 0.2% of addresses were duplicated in the 2011 Census because of more than one household at an address. The records were removed for this analysis to ensure we only used 2011 Census responses for an address that can be linked to the Valuation Office Agency (VOA) data.

For this analysis, a unique property reference number (UPRN) was assigned to households on the 2011 Census by using the method described in [the ONS working paper series no 17 - Using data science for the address matching service](#). Where we could not assign a UPRN, these cases were removed for this research. Where there were two or more records with the same UPRN, the record with the highest confidence score was selected. In situations where the highest confidence score was given to multiple UPRNs, all records were removed.

For this research, we used 2011 Census data without imputed values to ensure we only compared census responses with administrative data.

Valuation Office Agency property characteristics data

The VOA is an executive agency of Her Majesty’s Revenue and Customs (HMRC). It is responsible for banding properties for Council Tax.

The VOA property characteristics data do not contain information about individuals or households. The information that the VOA collects and holds about domestic and residential properties supports statutory functions for valuation and maintenance. For this research, we have used VOA property type and VOA dwelling code.

The VOA holds information on addresses, and it is not currently possible to clearly identify multiple households at an address from administrative data alone. If a property is empty, information would still be captured by the VOA but would not have been recorded by the 2011 Census as it would be classed as a non-response.

For this analysis, a unique property reference number (UPRN) was mapped to the VOA’s unique address reference number (UARN) for each address using the cross-reference table on AddressBase. Where we could not assign a UPRN or where there were cases of two or more records with the same UPRN, these cases were removed for this research. Additionally, properties built after 2011 were removed to allow better comparison with the 2011 Census.

[Research](#) into the quality of the VOA dataset as a source for statistics has been published.

10 . Related publications

[Admin-based statistics for property type, feasibility research: England and Wales](#)

Methodology | Last revised 21 February 2020

Early research demonstrating the potential of Valuation Office Agency data to provide detailed information on property type, how it compares to 2011 Census accommodation type, and a request for feedback on the usefulness of these statistics.

[Administrative Data Census Research Outputs: Housing Characteristics](#)

Webpage | Last updated 18 January 2021

These outputs are part of our ongoing research into the possibility of combining administrative data and surveys to produce statistics for census-type information about housing characteristics as part of an Administrative Data Census.

[Valuation Office Agency property attribute data: quality assurance of administrative data used in Census 2021](#)

Methodology | Last revised 31 July 2020

Summary of the quality assurance undertaken on administrative data for Valuation Office Agency property attribute data used in Census 2021.

[Data source overviews: Valuation Office Agency data](#)

Webpage | Last revised 27 June 2017

Data source overviews intended to provide a high-level view of new data sources being evaluated by the Census 2021 Transformation Programme.

[Admin-based statistics for property floor space, feasibility research: England and Wales](#)

Methodology | Last revised 21 February 2020

Early research demonstrating measures of property size (floor space) from Valuation Office Agency data, which statistics can be generated from these data, what they can and cannot be used for, and a request for feedback on their usefulness.

[Estimating the number of rooms in Census 2021: an update on deriving an occupancy rating from Valuation Office Agency number of rooms](#)

Article | 18 January 2021

Exploration of differences in room occupancy ratings derived from Census 2011 and linked Valuation Office Agency number of rooms data. The research aims to provide further assurance for the replacement of the number of rooms question in Census 2021.