

2011 Census Statistics for England and Wales: March 2011 QMI

Contact:

census.customerservices@ons.
gsi.gov.uk
+44 (0) 1329 444972

Release date:
16 May 2013

Next release:
To be announced

Table of contents

1. [Methodology background](#)
2. [Executive summary](#)
3. [Output quality](#)
4. [About the output](#)
5. [How the output is created](#)
6. [Validation and quality assurance](#)
7. [Concepts and definitions](#)
8. [Other information](#)
9. [Sources for further information or advice](#)

1 . Methodology background

National Statistic	
Survey name	
Frequency	Decennial
How compiled	Census
Geographic coverage	England and Wales
Sample size	
Last revised	16 May 2013

2 . Executive summary

Every 10 years since 1801, apart from 1941, the nation has set aside one day for the census – an estimate of all people and households in England and Wales. It is the most complete source of information about the population that we have. Every effort is made to include everyone, and that is why the census is so important. It is the only survey that provides a detailed picture of the entire population, and is unique because it covers everyone at the same time and asks the same core questions everywhere. This makes it easy to compare different parts of the country. However, no census is perfect and some people are inevitably missed.

The Office for National Statistics (ONS) therefore uses complex statistical techniques to adjust the census counts for those people missed by the census. The information the census provides allows central and local government, health authorities and many other organisations to target their resources more effectively and to plan housing, education, health and transport services for years to come. The latest census was held on Sunday 27 March 2011. This report describes the methodology used to produce the 2011 Census estimates and gives information about the quality of the census statistics.

This document contains the following sections:

- Output quality
- About the output
- How the output is created
- Validation and quality assurance
- Concepts and definitions
- Other information, relating to quality trade-offs and user needs
- Sources for further information or advice

3 . Output quality

This report provides a range of information that describes the quality of the output and details any points that should be noted when using the output.

We have developed [Guidelines for Measuring Statistical Quality](#); these are based upon the five European Statistical System (ESS) quality dimensions. This report addresses these quality dimensions and other important quality characteristics, which are:

- relevance
- timeliness and punctuality
- coherence and comparability
- accuracy
- output quality trade-offs
- assessment of user needs and perceptions
- accessibility and clarity

More information is provided about these quality dimensions in the following sections.

4 . About the output

Relevance

(The degree to which the statistical outputs meet users' needs.)

The census provides a once-in-a-decade opportunity to get an accurate, comprehensive and consistent picture of the most valuable resource of England and Wales – its population. The census provides the only source of directly comparable statistics for both small areas and minority population groups across England and Wales. It is used as a 10-yearly benchmark for our annual mid-year population estimates (MYEs), which are vital to central and local government for planning, monitoring and resource allocation. The 2011 Census of England and Wales was taken on 27 March 2011.

2011 Census statistics were released in phases over 2 years; more information on the content and timing of releases is available online in the [2011 Census Prospectus](#). These will delve further into the data and look at detailed population characteristics at local authority and ward levels as well as community characteristics at smaller geographical areas such as output areas.

The main users of census data include central and local government, the health sector, business, the education and academic community and members of the public. Important uses of census data include:

- allocating financial resources from central government
- targeting investment and aiding investment decisions
- planning and monitoring social and geographical change
- policymaking and monitoring
- academic and market research

Extensive consultation was undertaken with users of census data around the design and development of the 2011 Census questionnaire, the operation of the census, the statistical processes and the statistical output. More information about the consultations carried out can be found in the Assessment of user needs and perceptions section of this report.

We carefully evaluated all the suggestions submitted by users. The changes made for 2011 were those identified as being most needed by the major users of census information or those that would result in more reliable and accurate data. As a result of these consultations, new questions were developed and some existing questions were redeveloped so that more user needs are met.

Consultation on the content of the census has always resulted in much larger demand from users for questions than can feasibly be met, and the 2011 Census was no different. To meet the demand from users for census questions would have required over six pages of questions per person; when finalised, the 2011 Census questionnaire contained four pages of questions per person. Some questions were not included because the case made was not as strong as for other topics or questions; some were not included because question-testing found they were not acceptable to the public or testing resulted in an unacceptable drop in response rate (for example, income) and some were not included because testing found the results were not reliable (for example, sexual identity).

We believe that the 2011 Census questionnaire and census operational arrangements achieved a reasonable balance between the demands from users of census information, the burden on the public, and the concerns of the public in respect of the privacy of their information.

More details of the changes made for 2011 are given in the Comparability and coherence section of this report.

Timeliness and punctuality

(Timeliness refers to the lapse of time between publication and the period to which the data refer. Punctuality refers to the gap between planned and actual publication dates.)

The breadth and depth of census statistics means that the 2011 Census estimates were released in stages. The timetable was planned around user need and our aim was to ensure that statistics are released as soon as they are ready.

The first 2011 Census estimates were published on 16 July 2012 – 16 months after census day, 27 March 2011. This first release included usually resident population estimates for England and Wales at regional and local authority level by age and sex, and estimates of occupied households. These statistics were required for use in resource allocation and planning procedures undertaken by the Department for Communities and Local Government for local authorities in England and by government departments such as the Department of Health for public health.

The time lag between census day and 16 July reflected the time needed to carry out the Census Coverage Survey (CCS), to process the large volumes of census questionnaires (more than 20 million), to carry out complex statistical processes to produce population estimates adjusted for under- and over-coverage, and to fully quality assure the estimates. This has resulted in a consistent and complete set of census outputs that improved the quality and usefulness of the 2011 Census for users, but took slightly longer to produce than simply outputting the results without the benefit of statistical estimation and quality assurance.

On 22 October 2012, estimates of the number of residents of England and Wales who have a second address elsewhere by broad age group, sex, reason for the second address (working, holiday or other reason) and local authority were published.

On 23 November 2012, the 2011 Census population and household estimates for small areas in England and Wales were published. These included estimates for output areas (OAs), Lower Layer Super Output Areas (LSOAs), Middle Layer Super Output Areas (MSOAs) and wards (electoral divisions in Wales).

On 11 December 2012, Key Statistics tables for local authorities in England and Wales, and unitary authorities in Wales, were published. Tables that come under the Key Statistics banner provide information derived from more than one variable on the census questionnaire, and allow comparison across different areas. An example of a Key Statistics table is Tenure (owning a property outright or with a mortgage or renting – social or private) and household type (detached, semi-detached, terraced) by local authority.

On 30 January 2013, Quick Statistics and Key Statistics were published for more geographies including the output area hierarchies, administrative wards, Westminster Parliamentary Constituencies and civil parishes.

On 19 February 2013, Key Statistics for National Parks, and the Key and Quick Statistics for postcode sectors, health areas and Welsh government devolved constituencies were published.

On 26 March 2013, Quick Statistics on national identity, passports held, country of birth, and approximated social grade were published. Information for non-UK born short-term residents (economic activity, country of birth, passports held, and sex by age (5-year age group and single year of age) were also included in this release.

On 16 May 2013, the first set of Detailed Characteristics tables on migration, ethnicity, national identity, language, religion, passports, country of birth, health and unpaid care were released at local authority level.

Census outputs were published within the scheduled publication windows and in particular the population estimates in July 2013 were published in time to feed into the resource allocation to local authorities in financial year ending 2013. As well as the many policy-related purposes that census data are used for, there are many research uses for which the statistics will have continuing value for many years after the census.

Statistics on people with second addresses were published earlier than planned as the information was ready to publish and there was a user need. This information helps central and local government better understand the total number of people that may require services in their areas.

Although the statistical processing and quality assurance of the data is complete, work remained to be done to tabulate and check the results and ensure that the protection of personal information is guaranteed when detailed information is prepared for small geographic areas. The time lag between releases reflects the amount of time required to complete these tasks.

More information on the content and timetable of releases is available online in the [2011 Census Prospectus](#).

For more details on related releases, the [GOV.UK release calendar](#) provides 12 months' advance notice of release dates. If there are any changes to the pre-announced release schedule, public attention will be drawn to the change, and the reason for the change will be explained fully, as set out in the [Code of Practice for Official Statistics](#).

5 . How the output is created

An important objective of the 2011 Census is to provide high quality estimates that are required by main users on a consistent and comparable basis for small areas and small population groups.

Around 25 million pre-addressed questionnaires were posted out to all households using a specially developed national address register. Special enumerators delivered questionnaires by hand to people living in residential care homes, hospitals, hostels, boarding schools, university halls of residence, mobile home parks, marinas, military bases and other communal establishments.

Householders were given the choice to submit their answers to census questions online or by post. The questionnaires were electronically tracked allowing us to count the number of postal returns received, and importantly, to identify addresses for which a completed questionnaire had not yet been returned. This information was then used to deploy a team of around 30,000 collectors to focus the follow-up procedures on households and communal establishments who had not returned a completed questionnaire in those areas where response was lower.

Every effort was made to ensure everyone was counted in the 2011 Census, however, no census is perfect and some people are inevitably missed. This undercount does not occur uniformly across all geographical areas or across other sub-groups of the population such as age-sex groups.

To fill this gap, we developed a [Coverage Assessment and Adjustment Process](#) (CAA), which built on the [2001 One Number Census](#) (ONC) approach.

The methodology involved the use of standard statistical techniques, similar to those used by many other countries, for measuring the level of undercount in the census and providing an assessment of the characteristics of individuals and households missed. We then used this information to adjust the 2011 Census counts to include estimates of people and households not counted. This methodology was researched and developed over a number of years in consultation with academics, statisticians, demographers and users of census data. An [Independent Review of Coverage Assessment, Adjustment and Quality Assurance](#) methodology took place in 2011. The review team stated that, “the further procedures for quality assurance and adjustment significantly strengthen ONS's strategy for successful population estimation”.

The main stages of the method used to produce 2011 Census estimates can be summarised as follows:

- 2011 Census field work was carried out to enumerate the population
- data from census questionnaires was scanned, captured, coded and cleaned
- imputation techniques were applied to the cleaned data to correct for inconsistencies in answers and missing data
- the CCS was undertaken independently to establish the coverage of the census
- the CAA process was carried out to adjust the census data using the results of the Census Coverage Survey (CCS)
- the census estimates were quality assured to ensure they were the best they could be

Data capture, cleaning and coding

Questionnaire processing began by scanning the forms and capturing the data using optical character recognition. At the data capture stage, complex coding was used to assign numerical values to written text and tick box responses. This involved the use of coding rules and standardised national coding frames, such as Standard Industrial Classification 2007: SIC 2007 and [Standard Occupational Classification 2010](#): SOC 2010, which allow data to be easily compared between different sources.

The data were then loaded into a database and validated to ensure that the values for each question were within the range specified in the relevant coding frame. Next, duplicate responses were removed. These occurred when a household submitted more than one questionnaire, for example, both on paper and by internet, or recorded the same person more than once. Invalid responses were also removed, for example, dust on the questionnaires may have been captured as responses, respondents may have crossed through pages that did not apply to their household with the lines being captured as responses, or respondents may have accidentally skipped pages, completing their response over two different person records.

More information about this stage of processing is in [Data Capture, Coding and Cleaning for the 2011 Census](#).

Edit and imputation

As with any self-completion questionnaires, respondents to the census sometimes made mistakes when recording their answers. This resulted in missing data or invalid responses because they were inconsistent with other values on the questionnaire, for example, where a person gave their age as 5 and said they had a university degree. These mistakes could be unintentional, for example, where a respondent missed a question or thought they could tick more than one option, or intentional where a respondent either did not know the answer or did not want to provide the answer. If these mistakes were left in the data, the census statistics would look obviously wrong, damaging trust in this valuable and important dataset. We therefore developed an edit and imputation strategy to correct inconsistencies and estimate missing data whilst preserving the relationships between census characteristics. More information about this methodology is available in the [2011 Census Item Edit and Imputation Process](#). After item editing and imputation, all of the returned questionnaire records were complete and consistent. This stage of processing did not impute missing people; that was the purpose of the Coverage Assessment and Adjustment process described later in this section.

Census Coverage Survey

The purpose of the CCS was to improve the accuracy of census results by estimating the number and characteristics of people missed by the census. It was an independent voluntary survey involving the re-enumeration of all households and individuals in a sample of postcodes. A representative sample of 1.5% of all postcodes in England and Wales, covering 335,000 households, was included in the CCS. The response rate for the CCS was 90%, which is very high for a voluntary questionnaire.

Coverage assessment and adjustment

The aim of this methodology was to identify and adjust for the number of people and households not counted in the 2011 Census, those counted more than once, and those counted in the wrong place. It involved a number of stages:

- the CCS records were matched with those from the 2011 Census using a combination of automated and clerical matching
- the matched census and CCS data were used within a Dual System Estimation (DSE) technique to estimate the number of people and households missed by both the census and CCS
- the 2011 Census database was searched for duplicates and the CCS was used to estimate the level of overcount (those counted more than once) in the census
- populations for each local authority by age and sex were then estimated, balancing over- and under-estimates, using a combination of statistical regression and small area estimation techniques
- households and people estimated to have been missed by the census were then imputed into the census database

More details about the methodology can be found on the [2011 Census Quality and Methods](#) pages.

Quality assurance (QA) procedures were built into all stages of the CAA process and the 2011 Census estimates were subject to a rigorous QA process. This followed an agreed strategy, which had been the subject of wide consultation with census users. Further information about the QA process can be found on the 2011 Census Quality and Methods pages.

6 . Validation and quality assurance

Accuracy

Accuracy is the degree of closeness between an estimate and the true value.

Sampling error

Although the Coverage Assessment and Adjustment Process (CAA) methodology estimated and adjusted the census counts for those who did not respond to the census, estimates of the population were effectively based on a sample and are therefore subject to sampling error. As with any sample, different people would be selected if the sample was randomly drawn again and slightly different estimates would be produced based on this different sample. The spread of these estimates is known as the sampling variability. Confidence intervals are used to present the sampling variability.

A 95% confidence interval is a range within which the true population parameter would fall for 95% of all possible samples that could have been selected. It is a standard way of expressing the statistical accuracy of a survey-based estimate. If an estimate has a large error level, the corresponding confidence interval will be very wide. For England and Wales as a whole, the national population estimate had a 95% confidence interval of plus or minus 0.15%, suggesting that the true population count is expected to be within plus or minus 83,000 of the census estimate. [Confidence intervals for the 2011 Census](#) are available:

- by 5-year age groups by sex and sex ratios for each local authority in England and Wales
- for ethnicity and activity last week and tenure by census estimation area

The sampling error associated with the 2011 Census estimates is mainly dependent on the Census Coverage Survey (CCS) sample size, the size of the population, the census response rate, the CCS response rate and the degree of similarity of the population the error level relates to. At a national level, the overall error will be smaller than the error associated with a local authority, particularly one that has a low response rate or an area that has a diverse population.

Sample sizes do vary between local authorities and age-sex groups and therefore some error levels may be smaller or larger than average.

For similar reasons, the census estimates for smaller geographies will have more error associated with them. The methods used to produce the estimates were designed to achieve the highest quality at local authority level. This priority was agreed with the main users of the data. Although of high quality in absolute terms, the population estimates for small areas will be lower quality than those for local authorities. In England and Wales, the census was the best available source of statistics on the population of small areas in March 2011. There was no other reliable source of such detailed information available — that is partly why a census is required. Although we focused on the accuracy of local authority level census estimates, some quality checks of data for smaller geographic areas (Middle Layer Super Output Area (MSOA) and Lower Layer Super Output Area (LSOA) were carried out at lower levels, but these were not as extensive because there is no reliable source of small area statistics with which to compare the census.

Sampling error was minimised in the 2011 Census in several ways. The census fieldwork was designed to maximise overall response and minimise differences in response rates in specific areas and among particular population sub-groups. This was done using an up-to-date address register, which was developed together with a questionnaire tracking system to monitor return rates in real time. This information was used to target field staff to areas with lower response rates with the aim of reducing variability of response between areas and improve response in the lowest responding areas. In addition, the CCS was designed so that the sample was large enough to ensure that the accuracy of the estimates met quality targets, was representative of areas across England and Wales, and took into account the characteristics of areas that were hard to enumerate.

The 2011 CCS successfully achieved over 300,000 interviews and from it, it was therefore estimated that 6.1% of the population was missed by the 2011 Census. For more information see [2011 Census Coverage Survey Summary](#).

Non-sampling errors

Non-sampling error is the difference between an estimated value and the true value, which is not due to sampling variation. In the case of the census, non-sampling error can occur in most parts of the data collection and production process and can arise from four main sources:

- coverage error
- non-response error
- measurement error
- processing error

Coverage error arises from an inability to sample the entire population. Undercoverage would bias the results and make them less reliable. Reasons for undercoverage include non-return of questionnaires, and households not receiving a questionnaire because their household's address was missed by the address register. Undercoverage was adjusted for through the CAA process. An evaluation of the [Address Register](#) has been published.

Overcoverage can also occur because:

- duplicate returns were received from the same household
- duplicate returns were received from one individual (for example, a student is counted at their term-time address and also counted at their home address by their parents)
- an individual was counted in the wrong location (for example, a student is counted by their parents at their home address, but missed at their term-time address)
- errors were made by the individual, census collector or processing system (for example, people who are not usual residents of England and Wales, a baby born after census day, or someone who died before census day were incorrectly included)

Several processes were developed to correct for such overcount. More information is available in [Overcount estimation and adjustment](#).

Non-response is a potential source of error in a census. It is the error that occurs from failing to obtain some or all of the information from a member of the population. Such errors can contribute to the bias of estimates if non-respondents differ from respondents in their characteristics.

There are two sources of non-response error in the census; person or household non-response and item non-response. Person non-response error occurs because an individual does not respond to the census and household non-response occurs when an entire household fails to respond.

The 2011 Census person response rate is calculated as the number of usual residents whose individual details were completed on a returned questionnaire, divided by the estimate of the number of usual residents. For 2011, person response rates for England and Wales were estimated to be 94% and the household response rate was estimated to be 95%. Response rates varied across geographic areas, age and population groups. Details of person and household response rates for some main census variables, for example, by marital status and ethnic group are available in [Response rates in the 2011 Census](#).

Item non-response refers to missing or inconsistent values associated with a particular question, or set of questions, in an otherwise complete census questionnaire. Missing values occur typically when a respondent does not know or refuses to answer a particular question. Inconsistent values occur when responses to two or more questions are incompatible.

For instance, the data from one part of the questionnaire might indicate that the respondent is 5 years old, and the data from another part of the questionnaire might indicate that the respondent is in full-time employment. Item non-response can lead to bias in estimates derived from the data. The item non-response rate is the percentage of the measured population that had an invalid value for that item. It is calculated by dividing the total number of invalid responses for an item by the total number of persons who were required to answer that item.

Item imputation was applied to the census data to compensate for such bias by generating an estimated value where the answers to census questions were missing or inconsistent. By using actual data from other respondents with similar characteristics, the imputation process served to estimate and reflect accurately the distributional properties of a complete and consistent dataset. The item imputation rate is the percentage of the measured population whose values have been changed by the imputation process. It is calculated by dividing the total number of imputed responses by the total number of persons who were required to answer that item.

The difference between the item non-response rate and the item imputation rate is the item inconsistency rate. This is the percentage of responses that were replaced due to failing the edit rules, such that:

Total imputation equals non-responses plus inconsistencies

[Item non-response rates and item imputation rates](#) are available.

Measurement error is the error that occurs from failing to collect the correct information from respondents. Sources of measurement error in a census include a poorly designed questionnaire, errors by field staff or errors made by the respondent. Not all these errors can be measured.

In the summer of 2011, we carried out the [Census Quality Survey \(CQS\)](#), which was a small-sample voluntary survey to evaluate the information people provided on their census questionnaires. A team of interviewers visited selected households to ask census questions again in a face-to-face interview. The aim was to assess people's understanding of the questions and measure the accuracy of information collected in the census for all household members.

Agreement rates for some main questions are available on the Census Quality Survey pages. More results from the CQS were published with [future releases of 2011 Census data](#).

Processing error can be introduced by processes applied to the data before the final estimates are produced. It includes errors in geographical assignment, data capture, coding, data load, and editing of the data as well as in the CAA process described previously. It is not possible to calculate processing error exactly; however, various measures were taken during each process, which can be used as estimates of processing quality.

A total of 24 million census questionnaires were processed between March 2011 and November 2011. The information given on these had to be captured, converted into coded data, and cleaned so that the outputs produced would be of the highest possible quality and completeness. Targets were set to ensure that the captured and coded data were of sufficiently high quality. The minimum required level of accuracy for capture and coding varied by field type depending on the complexity of the data in the field. All of the targets set for data capture and coding accuracy were exceeded for the 2011 Census. More details about the data capture, coding and cleaning of census data are available in [Data Capture, Coding and Cleaning for the 2011 Census Data](#).

Another possible source of processing error is that introduced by the need to balance the level of detail that the analysis of the census data allows against protection of the confidentiality of the individual, which has always been paramount. In order to ensure confidentiality for 2011, statistical disclosure control (protecting the attributes of an individual) was applied to the data.

Targeted record swapping was used, which involves swapping a percentage of household records between geographic areas. While all households had a chance of being swapped, the swapping was targeted towards individuals and households with unique or rare characteristics. By targeting the swapping in this way the protection could be achieved by swapping a lower number of households than if the swapping had been done entirely randomly. Most swapping was done at MSOA level or below.

A similar targeted approach was used to protect residents in communal establishments, but individuals rather than households were swapped. As a result of using a method that targeted records where the risk of disclosure was greatest (mainly people and households with unusual characteristics who could be recognised in small areas), analyses based on larger numbers will not be greatly affected by disclosure control, however, the impact will be greater on smaller areas.

Due to the need to protect confidentiality of individuals, we do not publish swapping rates. For more information see [Statistical Disclosure Control](#) web pages.

A number of steps were taken to maximise response rates and reduce bias and errors in the census, including:

- the census questionnaire was well-designed and extensively tested
- the field operation was managed to maximise response rates and reduce variability using the address register and questionnaire tracking system
- nationwide and local publicity campaigns took place to explain the purpose and value of the 2011 Census, encourage householders to return completed questionnaires, and give the public assurances about confidentiality and data security
- we worked closely with local authorities and community groups to encourage participation in the census by all
- help was provided to the public via a 2011 Census website and telephone help-line
- questionnaires submitted online were automatically validated
- data captured was checked for validity and cleaned
- edit and imputation techniques were used to estimate missing data and correct for inconsistencies

Bias can be introduced into estimates if the assumptions on which a methodology is based are not met; for example, the Dual System Estimation (DSE) method relies on the independence between the census and Census Coverage Survey (CCS). The Coverage Assessment and Adjustment (CAA) methodology was designed to assess and address forms of bias that may have resulted from any violation in the underpinning assumptions. More information about the assessments and adjustments carried out are described in [The 2011 the 2011 Census Coverage Assessment and Adjustment Process](#).

Comparability and coherence

(Comparability is the degree to which data can be compared over time and domain, for example, at geographic level. Coherence is the degree to which data derived from different sources or methods, but refer to the same topic, are similar.)

Census information is available for the last 210 years – every 10 years since 1801, except for 1941 when no census was held due to the Second World War. While a census gives an excellent snapshot of the country at the time, changes in definitions, questions, categories used to present results, and geographical boundaries mean that direct comparisons between one census and another does not necessarily give the best estimate of broad population change. This is particularly true for comparisons of population estimates between the 2011 and 2001 Censuses as the 2001 Census is known to have underestimated the population. More information about the reasons for this is given in the [Local Authority Population Studies: Full report](#).

Further studies led to the Office for National Statistics (ONS) adjusting the 2001 [Mid-year Population Estimates](#) (MYEs) for England and Wales by 275,000 to take account of this underestimation, however, the 2001 Census database itself was not revised. For this reason, the best source to compare with the 2011 Census results to calculate population change over the decade at national and local authority level is the 2001 ONS MYEs.

The census in England and Wales gathers data on the population at the time of the census – in the case of 2011 this was 27 March. The annual MYEs provide updated estimates of the population as of 30 June between census years by ageing the previous year's population by 1 year (1 year and 3 months in the first year after the census) and accounting for births, deaths and migration estimated to have occurred during the year.

During census years the MYEs are calculated by ageing the population by the period of time between the census and 30 June and using information on the components of population change during that period to update the population base. On 25 September 2012, we released [mid-2011 population estimates for England and Wales](#), which were based on the results of the 2011 Census, updated to the mid-year reference date. The 2011 Census population estimates will also be used to rebase the MYEs going back to mid-2002 to ensure a consistent time series. These [revised estimates for mid-2002 to mid-2010](#) were published in December 2012 for England and Wales and March 2013 at subnational level within England and Wales. The [revised back-series for the UK as a whole](#) were published later in 2013.

Mid-2011 population estimates for Middle and Lower Layer Super Output Areas were published in spring 2013. Estimates for [wards](#), [Parliamentary Constituencies](#) and National Parks followed the publication of these estimates. Population estimates for smaller geographies for mid-2002 to mid-2010 were also revised to take account of the results of the 2011 Census and to ensure that they remain consistent with population estimates for local authority areas. The [revised estimates](#) were published later in 2013. More information about the methodology for producing the MYEs is published in [Population Estimates Methodology Guides](#).

For the 2011 Census, comparability has been retained with the 2001 Census and other ONS population statistics where possible, as this was an important design principle in the development of the questionnaire and the processing of the statistics. It was also a strong message gathered from users during the output consultation. For most topics, 2011 Census outputs will be comparable with those from 2001, so analysis of trends over time will be possible.

However, some changes in concepts and definitions were necessary to take into account societal changes in the decade, for example, the introduction of same-sex civil partnerships, and to improve the quality of the data collected and harmonisation with other data sources. The main differences between the 2001 and 2011 Censuses are summarised in the following section.

Changes in population definitions

The 2011 questionnaire provided explicit guidance about who should be included as a usual resident, ensuring that the England and Wales census used the same definition of usual residence as required by the United Nations Economic Commission for Europe (UNECE) regulations and as used in the ONS mid-year estimates (MYEs). This will deliver closer comparability between the census results, the MYEs and population estimates from other countries.

The “household” definition has been improved to make it easier to understand, more relevant to current living arrangements and to ensure consistency with the UNECE definition.

Two new questions about second residence were introduced in 2011. Together with the information on usual address, responses to these new questions will provide users with more information about complex living arrangements, for example, people who live away from home part of the week, and will help reconcile the census estimates with the MYEs. The information will be particularly useful for housing and transport planning as required by local authorities who will want to know the estimated number of people who stay within their area and use local services during the week but who have a usual residence elsewhere.

The 2011 Census is the first census of England and Wales to capture information on short-term residents through the inclusion of new questions on date of entry into the UK for in-migrants and their intention to stay. This addition is a direct result of user consultation and reflects the changing needs of census users. The definition used in the UK for short-term residents is coherent with ONS’s Short-Term Migration estimates for England and Wales; however, there are some subtle differences (see the Concepts and definitions section).

The [enumeration base](#) for 2011, used to determine who should complete a full census return, included usual residents and, unlike in 2001, non-UK short-term residents. However, it did not include visitors to the UK, who were only required to give basic information (name, date of birth and usual address and country of usual residence).

In the 2001 Census, filters were added to the questionnaire, which meant that labour market and travel to work questions were only asked of people aged 16 to 74. In 2011, the upper age limit was removed so that everyone aged 16 and over was asked the questions. This makes the census more complete and more representative of the England and Wales population. It also reflects users' requirements for information and provides for the inclusion of older respondents. However, for reasons of statistical disclosure control and for 2001 comparability, the second release of 2011 Census statistics uses the 16 to 74 age group. Subsequent releases will have age breakdowns for the age 16 and over population.

Changes in geographic boundaries

Maintaining stability in small area geography to allow comparisons over time was critical for the 2011 Census. Changes to 2001 Output Areas (OAs) and Super Output Areas (SOAs) were necessary, however, in areas where the 2011 Census indicated there has been significant population change since 2001. The majority (97.4%) of the 2011 OAs remain unchanged meaning that they can be directly compared with 2001. Of the other OAs:

- almost 2% (1.8%) have been split into two or more OAs; for these, direct comparisons can be made between estimates for the single 2001 OA and the estimates of the two or more 2011 OAs aggregated together
- around 1% (0.6%) were merged with one or more other 2001 OAs so direct comparisons can be made between the estimates from the 2001 OAs, aggregated together, and the single 2011 OAs' estimates
- the remaining 0.1% have been redesigned mainly because of local authority boundary; these cannot easily be compared with an equivalent 2001 OA, and therefore like-for-like comparisons of 2001 and 2011 estimates in these instances are not possible

More information about the changes in small area geography between 2001 and 2011 are available in [Changes to Output Areas and Super Output Areas in England and Wales, 2001 to 2011](#). For the previous reasons, the best source to compare with the 2011 Census results to calculate population change over the decade at larger geographies of local authority and above is the 2001 ONS mid-year estimates (MYEs).

2011 Census estimates for output geographies are aggregations of whole OAs, best-fitted to the geographies that were current as at 31 December 2011. This is the method used to produce all 2011 Census and national statistics, so that statistics estimates produced on the same geography are consistent, comparable and non-disclosive. The only exception to this are the estimates for national parks, which are exact-fit, as best-fit estimates were considered to be inappropriate for this largely rural geography. [An overview of best-fitting](#) explains how 2011 Census estimates were built from output areas.

Changes to questions asked in the census

The most significant difference between the 2011 Census questionnaire and the 2001 version was in the topics covered. There were nine new topics: number of bedrooms, type of central heating, second address, month and year of arrival in the UK, intended length of stay in the UK, national identity, passports held, main language and visitor information.

More detail about the comparability between the 2011 and 2001 England and Wales Censuses is available on the [Comparability over time](#) web pages.

Changes to methodology

The Coverage Assessment and Adjustment process, the edit and imputation strategy and statistical disclosure control have all built on the approaches used in 2001. For more information see the methodology papers for each of these on the [2011 Census Quality and Methods](#) web pages.

UK comparability

[National Records of Scotland](#) (NRS) is responsible for disseminating 2011 Census statistics for Scotland and [Northern Ireland Statistics and Research Agency](#) (NISRA) is responsible for disseminating 2011 Census statistics for Northern Ireland. The UK population estimates are collated by ONS. The first release of UK population estimates was on 17 December 2012 and covered the UK and individual country estimates by 5-year age bands. Subsequent releases will include single year of age estimates for UK and the individual countries, plus 5-year age bands for regional and local authority areas across the UK. More information on the content and timetable of future UK releases is available online in the [2011 Census Prospectus](#).

Coherence with other data sources

Users often compare population estimates for individual local authorities to the numbers of people registered on other data sources, for example, administrative records. These other data sources were used extensively in the quality assurance of the 2011 Census estimates, however, comparisons between datasets should be treated with caution. These datasets were set up for specific administrative purposes so are not designed to measure the whole population.

There are definitional differences to the census in the data collected, differences in recording practices and data quality issues. Many administrative datasets will include people who are not “usually resident” in a local authority according to the census definition. For example, the GP Patient Register and National Insurance registers will include people who are living in the UK for less than 12 months, or whose family home is in another part of the UK. The census questionnaire was designed to explicitly identify such people and, through processing, they were excluded from the usually resident population estimate.

A paper summarising the strengths and limitations of each source in relation to these topics has been published in [Overview of Administrative Comparator Data Used in 2011 Census Quality Assurance](#) and comparator data for each local authority can be downloaded from the 2011 [Census quality assurance pack](#). An analysis of the differences between GP Patient Registers and the 2011 Census is published in [Comparison between 2011 Census estimates and the GP NHS Patient Register](#).

Coherence between 2011 Census statistics and other data sources accompany each release. These can be found on the [2011 Census Data page](#).

7 . Concepts and definitions

(Concepts and definitions describe the legislation governing the output, and a description of the classifications used in the output.)

The census in England and Wales is required by law under the [1920 Census Act](#), as amended by the [Census \(Amendment\) Act 2000](#) and the [Statistics and Registration Service Act 2007](#).

The definitions used for the 2011 Census are consistent with the standard UNECE recommended definitions and have also been adopted by Scotland, Northern Ireland and the rest of the European Union. For more information about definitions see the [2011 Census User Guide](#).

8 . Other information

Output quality trade-offs

(Trade-offs are the extent to which different dimensions of quality are balanced against each other.)

As mentioned in the Relevance section, there is a trade-off between what users want from a census and the quality of the resulting data. The inclusion of questions on topics that are sensitive or difficult to answer could have a significantly adverse effect on the census as a whole, particularly the level of response.

There is also a balance to be struck between user needs, accuracy and the timeliness of results. Census estimates which have not been subject to coverage assessment and adjustment, or detailed quality assurance processes could be produced more quickly. However, these would not produce the complete and consistent census estimates that users require, and would not accurately reflect the population.

There is a trade-off between relevance to users and consistency and comparability with previous censuses. While continuity between censuses is extremely valuable, the census must also adapt to reflect the changes in society.

Assessment of user needs and perceptions

(The processes for finding out about users and uses, and their views on the statistical products.)

The design and content of the 2011 Census has been shaped by three principal determinants:

- the demands and requirements of users of census statistics
- the evaluation of the 2001 Census
- the advice and guidance of international census agencies and organisations with experience of similar operations

The main users of census data, who were invited to take part in all user consultation, are described in the Relevance section of this report. For a description of the main uses of census data, see the [Census 2011 White Paper: Helping to Shape Tomorrow](#) and the [Census 2011 Outputs web page](#).

Consultations took place through a structure of formal advisory committees, topic-related working groups and public meetings, via media events such as ONS consultation and information papers, and the 2011 Census website. The consultation was undertaken around the design and development of the 2011 Census questionnaire, the operation of the census, statistical processes and the statistical output.

The Census 2011 consultation process began in 2003 by looking at lessons learnt from the 2001 Census. This process continued through to 2011, involving a variety of consultation methods to meet different census user needs, culminating in Output Roadshows run across England and Wales enabling users to provide feedback on the evolving census outputs design and content.

Details of the process undertaken to develop the content of the 2011 Census questionnaire for England and Wales are in the report [The 2011 Census: Final questionnaire content for England and Wales](#).

Further information on the [Census 2011 consultation process and results](#) is also available.

9 . Sources for further information or advice

Accessibility and clarity

(Accessibility is the ease with which users are able to access the data, also reflecting the format in which the data are available and the availability of supporting information. Clarity refers to the quality and sufficiency of the release details, illustrations and accompanying advice.)

Our recommended format for accessible content is a combination of HTML web pages for narrative, charts and graphs, with data being provided in usable formats such as CSV and Excel. The Office for National Statistics website also offers users the option to download the narrative in PDF format. In some instances other software may be used, or may be available on request. Available formats for content published on the ONS website but not produced by ONS, or referenced on the ONS website but stored elsewhere, may vary. For further information please refer to the contact details at the beginning of this report.

For information regarding conditions of access to data, please refer to the following links:

- [Terms and conditions](#) (for data on the website)
- [Copyright and reuse of published data](#)
- [Pre-release access](#) (including conditions of access)
- [Access to microdata](#) via the Virtual Microdata Laboratory
- [Accessibility](#)

2011 Census statistics are published on the main Office for National Statistics (ONS) website, as well as the Nomis website. All standard outputs are free under the [Open Government Licence](#). Users will be able to find and select 2011 Census outputs, and choose to download or view the statistics online. Supporting reference materials will accompany the statistics on the [2011 Census User Guide](#) pages.

Each release is accompanied by statistical bulletins, commentaries, metadata, look-up files, a glossary of terms, information about the variables and classifications used and data visualisations as appropriate. These supporting documents aid the clarity and understanding of the census data and ensure they are used appropriately.

The data was released in stages, and a series of specialist products followed. These include data on alternative population bases, microdata samples, flow data (origin and destination data) and statistics on small populations. Some of these later releases were available via the Virtual Microdata Laboratory (VML), which is hosted by ONS. This is available to approved researchers and other interested parties through a rigorous vetting procedure.

As well as developing the main ONS website to accommodate users' needs for accessible, online statistics for use and exploration, published census data will be directly accessible to third parties to power their own websites and applications, using an Application Programming Interface (API).

Further information about the [2011 Census](#), releases, frequently asked questions, a comparison between the UK and Europe censuses and other information is also available.