

# Population estimates by output areas, electoral, health and other geographies QMI

Quality and methodology information for national population estimates broken down into small geographical areas (Super Output Areas, health geographies, electoral wards, Parliamentary constituencies and National Parks) in England and Wales. Details the strengths and limitations of the data, methods used, and data uses and users.

Contact:  
Neil Park  
pop.info@ons.gov.uk  
+44 1329 444661


Release date:  
16 September 2021

Next release:  
To be announced

## Table of contents

1. [Output information](#)
2. [About this Quality and Methodology Information report](#)
3. [Important points](#)
4. [Quality summary](#)
5. [Quality characteristics of the population estimates data](#)
6. [Methods used to produce the population estimates data](#)
7. [Other information](#)

# 1 . Output information

<b>National Statistic, Experimental Statistics and supporting information depending on geography and age disaggregation</b>	
<b>Survey name</b>	Population estimates by output areas, electoral, health and other geographies
<b>Frequency</b>	Annual
<b>How compiled</b>	Based on third party data
<b>Geographic coverage</b>	England and Wales
<b>Last revised</b>	16 September 2021

## 2 . About this Quality and Methodology Information report

This quality and methodology report contains information on the quality characteristics of the data (including the five European Statistical System Dimensions of Quality) as well as the methods used to create it. The information in this report will help you to:

- understand the strengths and limitations of the data
- learn about existing uses and user of the data
- reduce the risk of misusing data
- help you to decide suitable uses for the data
- understand the methods used to create the data

## 3 . Important points

- Population estimates by output areas, electoral, health and other geographies (formerly known as small area population estimates (SAPE)) are the official source of population estimates between census years for Output Areas, Super Output Areas, clinical commissioning groups, Parliamentary constituencies, electoral wards and National Parks; for the purpose of this QMI, we will continue to refer to these estimates under the SAPE acronym.
- Output Areas and Super Output Areas are designed as a "building block" geography and so part of the purpose of this release is to allow users to build estimates for customised areas using official population estimates.
- The estimates use the census definition of people who are "usually resident" in the UK for 12 months, excluding short-term migrants and counting students at their term-time addresses.
- The estimates roll forward the population found by the previous census, one year at a time by using the ratio of change in consecutive GP Patient Register data and changes to special populations to account for population changes; estimates for small areas are constrained to local authority level estimates.
- SAPE have [National Statistics](#) status for: Middle layer Super Output Areas for five-year age groups by sex, Lower layer Super Output Areas for broad age groups by sex, and clinical commissioning groups.
- On 25 October 2018, [a revised set of SAPE for mid-2012 to mid-2016](#) was published; SAPE have been revised in light of revisions to the [mid-year population estimates, published in March 2018](#), that incorporated methodological improvements for local authority level estimates or immigration and emigration within England and Wales.
- The next set of estimates and revisions are not expected until the results of the Census 2021 are available.
- The mid-2020 estimates reflect the population in England and Wales after the first four months of the coronavirus (COVID-19) pandemic.
- The pandemic has impacted several data sources used in the regular production of population estimates but it is not possible to quantify the impacts on quality.

## 4 . Quality summary

This report relates to the set of small area population estimates (SAPE), which includes population estimates for Lower layer Super Output Areas (LSOA), Middle layer Super Output Areas (MSOA), electoral wards, Parliamentary constituencies, health geography (clinical commissioning groups (CCGs), NHS England (Region) and NHS England (Region, Local office) and National Parks. Estimates are available annually by age and sex, from mid-2001 onwards.

Estimates for LSOAs by broad ages and MSOAs by five-year age groups (quinary age) hold [National Statistics](#) status. Estimates at a greater level of disaggregation by age including quinary age for LSOAs and single year of age for both SOAs are supporting information only. However, we intend to conduct analysis to assess whether the LSOA (and OA-level estimates) by five-year age groups are of suitable quality for National Statistics status.

Electoral wards, Westminster Parliamentary constituencies and National Parks all hold [Experimental Statistics](#) status. Estimates for health geographies are aggregated directly from LSOAs and hold National Statistics status. From mid-2012 onwards, the SAPE can be broken down into two main types: the Super Output Area (SOA) estimates (both middle and lower layer), produced using a "Ratio Change" methodology and the estimates for other geographies produced using an "Output Area Apportionment" methodology. The rationale behind both of these approaches is to use information held in administrative datasets to update estimates of the population year-on-year, until results from the next census can be used to rebase the data.

Estimates for mid-2011 used an alternative methodology as they are based on the results of the 2011 Census, updated to account for the population change that occurred between census day (27 March 2011) and the mid-year point (30 June 2011). Revised census-based estimates for the intercensal period (mid-2002 to mid-2010) have been produced, to provide a consistent time series of population estimates. The original estimates for mid-2002 to mid-2010 were produced using a "Postcode Best-fit" methodology, which is no longer in use.

## Uses and users

### Assessment of user needs and perceptions

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

Local area users are consulted through the Central and Local Information Partnership (CLIP) Population Subgroup. The aim of the CLIP Population Subgroup is to improve communication between the Office for National Statistics (ONS) and some users of population statistics produced by the ONS. Further information on this user group is available on the [Knowledgehub](#) website.

Users are encouraged to email feedback and suggestions regarding the population estimates to the Population Estimates Unit Customer Services Team [at pop.info@ons.gov.uk](mailto:at.pop.info@ons.gov.uk). Following the start of the pandemic in March 2020 we have received and acted upon a range of stakeholder feedback. This has included a widespread user need for earlier access to the data (accomplished by moving the release date to September) and enquiries from users around making additional geographic breakdowns and time series available.

User views are sought specifically on population outputs including SAPE. A formal consultation on SAPE held in June and July 2014 helped to identify the large number of uses made of SAPE data, along with useful feedback on users' views on their quality. A [summary of the results](#) is available.

### Strengths and limitations

Main strengths:

- the estimates provide timely, official data on the size, age and sex of the population between censuses
- methods are held as consistent as possible; major changes are accompanied with a back-series allowing users to understand any changes
- these estimates are coherent with [Population estimates for UK, England and Wales, Scotland and Northern Ireland: mid-2020](#)

Main limitations:

- the estimates only cover the usually resident population and so do not include "daytime" populations or short-term visitors
- the ratio change method used to produce these estimates assumes that change in administrative data is a good proxy for change in the population of an area; while this assumption holds reasonably well for the majority of areas, administrative processes (such as list cleaning) can cause issues
- the methods used mean that the smallest geographical entities for which populations are available are output areas and estimates for many geographical areas are only available on a best fit basis

Information on the methods used to produce these components of population can be found in the methods guides:

[Methodology note on production of population estimates by output areas, electoral, health and other geographies, England and Wales: mid-2020.](#)

## Review of population estimates and projections

In May 2021, the Office for Statistics Regulation (OSR) published a [review into population estimates and projections](#). This highlighted that at the national level, the approach taken by the ONS is fit for purpose and is supported by expert advice from demography and academic partners but that at the subnational level, the accuracy of estimates is variable because of factors such as the size and mobility of the population in a given area.

The review also identified the need to keep methods current and responsive, especially at some lower levels where there is more variability in the data. We are already progressing work in this area as we consider how best to meet these recommendations, and we will report back to the OSR about our plans while continuing to gather feedback on them more broadly. We recognise the need for:

- development work to keep these sources current and responsive
- developing how we work and engage with users of our statistics, particularly at a local level
- planning how we move to Census 2021-based estimates, and take on board improvements from the population and migration statistics transformation programme

While the review did not explicitly discuss population estimates below local authority level, quality issues at the national and local authority levels will also have an impact on population estimates at all levels. Further, the methodology used to produce estimates for census, electoral and health geographies is less sophisticated than for higher geographies and there is likely to be pockets of uncertainty in these estimates.

## 5 . Quality characteristics of the population estimates data

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

We have developed [Quality in official statistics](#); 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
- accuracy and reliability
- output quality
- coherence and comparability
- concepts and definitions
- geography
- accessibility and clarity
- timeliness and punctuality
- why you can trust our data

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

## Relevance

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

The small area population estimates (SAPE) provide annual estimates of the population at mid-year (30 June), by age and sex.

Enquiries and requests show that the main users of SAPE are central government, local authorities, the health sector and academics. Main uses of the estimates can be classified into two broad groups:

- use of the absolute numbers (for example, for planning purposes)
- use of the estimates in a variety of rates and indicators

Discussions with users have identified that their priorities for the quality of the estimates are: that the data are timely; that they provide sufficient detail (by geography and age group); and that they are consistent with the Office for National Statistics' (ONS's) other population estimates.

From mid-2011, the estimates are published by single year of age and sex in order to meet user needs for more detailed information and allow users to create estimates for bespoke age categories. Revised estimates for mid-2002 to mid-2010 have also been published by single year of age and sex. Users are made aware of the limitations of estimates at this fine level of age detail.

To meet user needs for more detailed geographic information, mid-2011 estimates have been published for wards and Parliamentary constituencies on a census output area best-fit approach in line with the [ONS geography policy](#). Estimates for the underlying Census Output Areas are also available, which can be aggregated to produce Output Area best-fit estimates for any other geography. This gives users the flexibility of being able to obtain population estimates for locally-defined or bespoke areas. Mid-2002 to mid-2019 Output Area estimates have also been published. Users are made aware of the limitations of estimates at the Output Area level.

For consistency, all SAPE are constrained to the local authority mid-year population estimates. The revised mid-2002 to mid-2010 SOA estimates are constrained to the revised mid-2002 to mid-2010 local authority estimates.

## Accuracy

(The degree of closeness between an estimate and the true value.)

Population estimates for Lower layer Super Output Areas (LSOAs) by broad ages and Middle layer Super Output Areas (MSOAs) by five-year age groups (quinary age) and clinical commissioning groups (CCGs) hold National Statistics status. The remaining estimates for other geographies are published as [Experimental Statistics](#). This means that they have not yet demonstrated the quality criteria for National Statistics, but are being published to involve users in the development of the methodology and to help build quality at an early stage. More information on Experimental Statistics is provided in the UK Statistics Authority [Code of Practice for Statistics](#).

The reliability of the estimates is difficult to quantify due to the range of information used in the construction of the estimates. We quality assure the administrative data used for these statistics to ensure that they are suitable for this purpose. To gain further insight on data quality issues and the impact on statistics, please see the Quality Assurance of Administrative Data reports for the [Patient Register](#), [UK Armed Forces](#), [US Armed Forces and Prisoners](#).

Every 10 years, the national census provides a benchmark for population estimates. Other than comparisons with the census, there are no benchmark population counts against which the SAPE can be compared. Some elements of the population are more difficult to estimate, such as migrants (internal and international) and students. It is therefore expected that estimates for areas with a large proportion of more difficult to estimate elements of the population are likely to have larger errors developing over time.

Initial analysis has been completed comparing a set of "rolled forward" mid-2011 estimates produced using the ratio change methodology against the published census-based mid-2011 estimates. This analysis shows that the rolled forward estimate was within plus or minus 1% of the census-based estimate for mid-2011 in 5,222 LSOAs (15.6%) and within plus or minus 5% in 21,688 LSOAs (64.7%). This gives an initial indication of the likely impact of the errors that have occurred in the LSOA population estimates time series across the decade between censuses and was used to revise the mid-2002 to mid-2010 population estimates. More information can be found in [Small Area Population Estimates Evaluation: Report on Accuracy Compared to Results of the 2011 Census](#).

The accuracy of these estimates is likely to have been negatively impacted by the coronavirus (COVID-19) pandemic. Estimates for census, electoral and health geographies all make use of a ratio change method that assumes that changes in administrative data (specifically the GP patient register) are a good proxy for changes in the population. However, the pandemic disrupted both the movement of people and their ability to access services.

Evidence from several sources suggests that some population groups - in particular students - made internal migration moves in March and April 2020 in response to the coronavirus pandemic and the closure of university halls of residence. The evidence is not easily quantifiable but one example, from the [National Union of Students' Coronavirus and students survey](#) from April 2020 suggests that pre-pandemic, around 24% of students lived with parents or guardians but this increased to 54% during the early part of the pandemic. It is quite likely that a high proportion of moves made during the first wave of the pandemic were not accompanied by a timely updating of administrative data. Consequently, population estimates for some student areas may be slightly overestimated while populations elsewhere may be slightly underestimated.

## Output quality

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

## Revised mid-2002 to mid-2010 estimates

Research work undertaken prior to the publication of the revised mid-2002 to mid-2010 SAPE identified three possible methods to produce a backseries of population estimates.

A "full assessment method" using census and administrative data along with an individual consideration of each area, where required, would have resulted in more accurate estimates overall. However, the advantages of increased accuracy were weighed against the impact on timeliness — that is, how soon the estimates could be published. Here there was a trade-off between different aspects of the quality of the estimates.

The "distribution of the difference" method provided the best balance in the majority of small areas between producing a plausible backseries of population estimates for each individual area and using a relatively straightforward method to allow timely publication. The method was designed to identify the difference between the census-based and rolled-forward mid-2011 estimates for each Output Area (OA) and LSOA and distribute this difference across the backseries in order to remove any "jump" in the estimates between mid-2010 and mid-2011. As a consequence, the patterns of change identified in administrative data using the ratio change method may not be maintained in the revised mid-2002 to mid-2010 figures.

Limitations of this method are that it relies on making an assumption on how the difference between the two sets of estimates for mid-2011 has developed over time. This assumption will be particularly important for OAs or LSOAs where the 2011 Census estimates were very different from the "rolled-forward" estimates.

As the difference is distributed across the OA and LSOA backseries by age-sex cohort, an implicit assumption is also made that populations in mid-2011 would have been in an area in 2002 at a younger age (that is, a 19 year-old male in mid-2011 would have been in the same area in 2002 but aged 10 years). This was a particular issue in LSOAs with high student-aged populations. Constraining the LSOA estimates to the revised subnational mid-year estimates will have corrected for this to a certain degree: however, a minority of LSOAs show very small counts at younger ages as a result of this assumption. Care must be taken in interpreting age distributions for areas affected by this issue.

## Mid-2012 estimates

Mid-2012 SAPE rely on administrative data from the Patient Register to provide the detailed information about the age and sex distribution of the population at OA level. By contrast, the mid-2011 OA estimates are heavily based on the results of the 2011 Census with minor adjustments to account for population change during the period between census day and the mid-year point. In a minority of areas, where the census distribution is significantly different from that given by 2012 Patient Register data, there may be large differences between the mid-2011 and mid-2012 estimates for some OAs. The OA estimates are constrained to the LSOA totals, so in general estimates for higher geographies, which often contain whole LSOAs, do not show as much variation as that seen at the OA level.

The mid-2012 population estimates use administrative data to account for the special population (prisoners and armed forces) that is present in each small area. The mid-2011 estimates include 2011 Census estimates of the special population, which in some areas may differ significantly from those given by the administrative data sources for mid-2012. These definitional differences may create unexpected changes in population between mid-2011 and mid-2012 for a small minority of areas with large special populations.



## Mid-2013 estimates

Child Benefit data obtained from HM Revenue and Customs (HMRC) is one of the administrative sources used to indicate the level of population change in the 0-to-14-year-old age groups when calculating LSOA and MSOA mid-year population estimates. The Ratio Change method relies on the differences seen between the two extracts of administrative data being representative of any change in the underlying population rather than being the result of any policy or methodological changes. However, in January 2013, a new policy (the "High Income Child Benefit Charge") was introduced, which means that claimants are liable to repay some or all of their Child Benefit if they or their partner earn more than £50,000 per year. Alternatively, people earning above this threshold have the option to withdraw from the Child Benefit system.

The effect of these changes meant a large decrease in the number of children registered for Child Benefit between mid-2012 and mid-2013 (particularly in the very young age groups), which rendered the data unsuitable for use in estimating the population at small area level in mid-2013. This means that the SAPE for children aged 0 to 14 years have been calculated solely from Patient Register data in mid-2013, however, this is not thought to significantly impact the quality of the figures as the Patient Register has generally good coverage of people in these age-groups.

In some London local authorities the number of people included in Patient Register data in 2013 has decreased in a large number of LSOAs and MSOAs compared with 2012 data, which may be due to changes in administrative practices or may reflect genuine population change. The process of constraining LSOA and MSOA estimates to previously published local authority population estimates means that this pattern is not automatically reflected in the mid-year estimates. However, the constraining process has created some anomalous changes in a minority of LSOAs and MSOAs so some caution should be applied in interpreting estimates that show large percentage changes from the previous year.

## Forest Heath issue

SAPE for mid-2013 were affected by an error identified in the allocation of foreign armed forces in the mid-2013 local authority population estimates, an issue described in detail in the [quality and methodology information](#) report for those statistics. In total, the error affected population estimates for males and (to a lesser extent) females aged 18 to 59 years in 45 local authority areas, particularly Forest Heath in Suffolk.

As the SAPE are constrained to the local authority totals, these errors were carried forward into the figures for all small area geographies within (or containing) the affected areas. In the majority of areas the impact of the problem was negligible, however, some caution should be taken in interpreting figures for small areas known to have large numbers of foreign armed forces or located near foreign armed forces populations, particularly if these have changed significantly from the estimates for mid-2012.

When we identified the error we sought users' views on the preferred approach to correcting the estimates. We did this through direct contact with the most affected users, through Local Insight Reference Panel events, through posts on Stats Usernet and through publicising the error in our published outputs.

Taking into account the responses we received, we decided to:

- retain the previously published 2013 SAPE as the current estimates until autumn 2015
- provide information on the scale of the error for different ages within affected LSOAs
- publish corrected 2013 estimates alongside the 2014 estimates being released in November 2015; the current 2013 estimates remain on our website but are superseded by that new release

This approach is consistent with what was done for the local authority level mid-year estimates, where corrected 2013 estimates were released with 2014 estimates in June 2015.

The supporting information on the scale of the error for different ages within affected LSOAs is published as [Table 003757](#).

## Mid-2014 to mid-2016 estimates

In some local authorities the number of people included in Patient Register data in for example, 2016 has increased or decreased in a large number of LSOAs and MSOAs compared with 2015 data, which may be due to changes in administrative practices or may reflect genuine population change. The process of constraining LSOA and MSA estimates to previously published local authority population estimates means that this pattern is not automatically reflected in the mid-year estimates.

Following the publication of the 2014 LSOA estimates, concerns were raised about the quality of estimates (since 2011) for the population aged 90 years and over in one LSOA area of Denbighshire. We have addressed this issue in the revised SAPE back series (2012 to 2016) produced in 2018 following the introduction of methodological improvements to local authority population estimates.

## Revised estimates for mid-2012 to mid-2016

Estimates for mid-2012 to mid-2016 were revised in October 2018 to be consistent with the [revised local authority level estimates published in March 2018](#).

As part of this process the data were re-processed for mid-2012 and mid-2013 to be consistent with the methods and data used for the original production of mid-2014 to mid-2016. Therefore, the estimates for mid-2012 and mid-2013 have been processed using the system originally introduced for the mid-2014 estimates and estimates for mid-2012 have been processed using Patient Register data ratio change only (originally this used a combination of Patient Register and Child Benefit data).

The issues with the Denbighshire LSOA have also been addressed in the revised back-series. This has led to a minor discontinuity between estimates for mid-2011 and mid-2012 but more realistic estimates for mid-2012 onwards.

In a very few circumstances the underlying Patient Register data have been modified to prevent estimates for age and sex groups in some areas negatively impacting on the estimates of other areas within MSOAs or within a local authority once estimates are constrained to MSA or local authority totals. Less than 10 of these changes (to LSOA quinary age or sex groups) have been made over the six years of the revised series.

An interactive map showing the impact of these changes at LSOA level by broad age is provided in [Small area population estimates in England and Wales: mid-2017](#).

## Mid-2017 to mid-2020 estimates

In some local authorities the number of people included in Patient Register data in, say, 2020 has increased or decreased in a large number of LSOAs and MSOAs compared with 2019 data, which may be because of changes in administrative practices or may reflect genuine population change. The process of constraining LSOA and MSA estimates to previously published local authority population estimates means that this pattern is not automatically reflected in the mid-year estimates.

In a very few circumstances, the underlying Patient Register data have been modified to prevent estimates for age and sex groups in some areas negatively impacting on the estimates of other areas within MSOAs or within a local authority once estimates are constrained to MSA or local authority totals.

## Coherence and comparability

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

The SAPE are consistent with the national and local authority mid-year estimates, whereby counts by age and sex all sum to higher geography estimates. Small cell adjustment, as part of the disclosure control in 2001 Census-published outputs, has been applied to the revised mid-2002 to mid-2003 estimates (which rely very much on census counts), in order to maintain that level of protection. Revised estimates for mid-2004 to mid-2010 have also been subject to disclosure control.

There are other producers of small area population estimates including a number of local authority producers. The estimates these local authorities produce are likely to differ from the ONS estimates as the data used by the ONS to produce our estimates are not generally available and some authorities may use local data sources, for example, information on housing completions and demolitions. Some analysis at ward and MSOA level has, however, been done comparing ONS estimates with local authority estimates produced by different methods and this has shown there to be broad consistency with all the estimates compared.

National Records of Scotland (NRS) and the Northern Ireland Statistics and Research Agency (NISRA) produce small area population estimates for [Scotland](#) and [Northern Ireland](#) respectively. These estimates are produced using different methodologies from those used by ONS to produce the SAPE for England and Wales.

The published LSOA and MSOA population estimates provide a consistent time series, based on a stable geography, which is not subject to boundary change between censuses. The estimates are also consistent with published estimates for other geographies, for example, national (England and Wales) and local authority mid-year estimates. The published estimates for other geographies can be produced for all years on the latest boundaries to ensure that a consistent time series is available. In general, estimates are provided on the latest boundaries available.

A number of revisions have been made to the SAPE in order to reflect revisions made to the local authority mid-year population estimates:

- 2007 - revisions made to the mid-2002 to mid-2004 LSOA and MSOA population estimates to reflect revisions made to the local authority mid-year estimates in August 2007 incorporating some improvements made to mid-2001 estimates
- 2010 - revisions made to the mid-2001 to mid-2008 LSOA and MSOA population estimates to reflect revisions made to the local authority mid-year estimates (for mid-2002 to mid-2008) in May 2010, plus a few improvements to the mid-2001 estimates in a small number of areas
- 2011 - revisions made to the mid-2002 to mid-2007 population estimates based on the postcode best-fit (PBF) methodology to reflect revisions made to the local authority mid-year estimates in May 2010
- 2013 - revisions to the estimates for mid-2002 to mid-2010 for all geographies to provide a consistent time series from mid-2002 to mid-2011, given the results of the 2011 Census; a number of corrections for known issues were also made during the production of these revised estimates
- 2015 - revisions to the estimates for mid-2013 to correct for the processing error mainly affecting the local authority of Forest Heath and the small area estimates within it
- 2018 - revisions to the estimates for mid-2012 to mid-2016 in light of revisions to local authority population estimates made in March 2018; further details can be found under other output quality

These revisions are in keeping with the [ONS revision policy for population statistics](#).

## Concepts and definitions

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

The mid-year population estimates are consistent with the standard UN definition for population estimates, which is based upon the concept of usual residence and includes people who reside in the country for at least 12 months, whatever their nationality. Visitors and short-term migrants (who enter or leave the UK for less than 12 months) are not included.

Members of Her Majesty's armed forces stationed in England and Wales are included at their place of residence but those stationed outside of England and Wales are excluded. Members of the US armed forces stationed in England and Wales are included.

Students are taken to be resident at their term-time address. For mid-2020 this is important to be aware of as there is evidence that a large number of students left their usual term-time address as student halls were closed during the first wave of the coronavirus pandemic. For the most part, students in mid-2020 will be estimated at their term-time address.

Prior to 2011, prisoners had been regarded as usually resident at an institution if they have served six months or more of a custodial sentence. However, from 2011 onwards, this definition has changed to those who have been sentenced to serve six months or more, which is consistent with the definition used in the 2011 Census.

The figures for the UK do not include the population of the Channel Islands or the Isle of Man.

For a minority of people the concept of usual residence is more difficult; homeless people would be included in the population estimates if counted in the previous census but could be allocated to the wrong area if subsequent moves were not captured by the patient registers.

Although usual residence is the recognised definition for population estimates, use of a single definitional base does not meet the needs of all users. The usually resident population does not always coincide with the number of persons to be found in an area at a particular time of day or year. The daytime populations of cities and the summer-time populations of holiday resorts, for example, will normally be larger than their usually resident populations. We have also developed and published [national estimates of short-term migrants](#) to supplement the mid-year population estimates. These estimates refer to the flows of short-term migrants to and from England and Wales for each year since mid-2004. Short-term migration estimates at local authority level are also available.

## Geography

SAPE are produced by the ONS for England and Wales, by Output Area, Lower and Middle layer Super Output Areas, electoral wards, Parliamentary constituencies, clinical commissioning groups (England only) and National Parks. SAPE estimates are published unrounded by single year of age and sex at the lowest levels of geography possible to allow users to aggregate estimates to bespoke geographical areas as needed. The [ONS geoportal](#) provides a range of lookup files that allow users to produce estimates for major towns and cities, built-up areas and historic wards. It also provides a [hierarchical representation of UK statistical geographies](#) a diagram showing how different geographies fit together, and a short description of each geography in the [Beginner's Guide to UK Geography](#).

Each year we publish estimates for the boundaries in place at the reference point and for the latest boundaries where these differ. While most boundaries remain similar from year to year there are annual changes to both CCG and ward boundaries. Additional geographic breakdowns are published on [Nomis](#).

SAPE include a number of geographies that would not usually be described as small areas, such as CCGs and National Parks, which are often larger than local authority areas. This is due to these estimates being created from smaller building blocks (either LSOA or OA-level estimates) as the areas do not align with local authority boundaries.

## Assessment of user needs and perceptions

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

User views are sought specifically on population outputs including SAPE. A formal consultation on SAPE held in June and July 2014 helped to identify the large number of uses made of SAPE data, along with useful feedback on users' views on its quality. A [summary of the results](#) is available.

Local area users are consulted through the Central and Local Information Partnership (CLIP) Population Subgroup. The aim of the CLIP Population Subgroup is to improve communication between ONS and some users of population statistics produced by ONS. Further information on this user group is available on the [Knowledgehub](#) website.

Users are encouraged to email feedback and suggestions regarding the population estimates to the Population Estimates Unit Customer Services Team at [pop.info@ons.gov.uk](mailto:pop.info@ons.gov.uk).

## 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.)

Due to the requirement for SAPE to be consistent with the local authority mid-year estimates, the publication of the LSOA and MSOA estimates and other geographies follows on from the publication of these local authority estimates. It is the aim that the SAPE should be published as soon as is practically possible after the local authority estimates.

From mid-2013, all SAPE (including LSOA and MSOA estimates along with Output Area-based estimates for other geographies) were published in a single release. This was to reduce the number of releases and make it easier for users to find the data. The mid-2019 and mid-2020 estimates were published approximately three months after the local authority-level mid-year estimates, that is, 15 months after the reference date of the estimates. Improvements have been made in order to reduce this time lag since the estimates were first published (mid-2005 estimates were originally published 28 months after the reference date).

When revisions have been made to the local authority mid-year estimates (and the back-series of estimates) as a result of methodological improvements, the SAPE have also been subject to revision. Where this situation has occurred, the time lag between the reference and publication dates of the estimates has been increased due to the additional work required to incorporate the revisions. The majority of revised SAPE estimates for mid-2002 to mid-2010 were published approximately seven months after the revised local authority mid-year estimates (National Park estimates were not published until 10 months after the local authority figures).

All SAPE have been published to time on pre-announced dates apart from a single slippage of one week.

For more details on related releases, a [release calendar](#) for government statistics provides 12 months' advance notice of release dates. In the unlikely event of a change to the pre-announced release schedule, public attention will be drawn to the change and the reasons for the change will be explained fully at the same time, as set out in the [Code of Practice for Statistics](#).

## Mid-2019 and mid-2020 estimates

For both the mid-2019 and mid-2020 releases the publication date has been shifted forwards by around two months from late October to early September. This accelerated production schedule reflects a higher level of interest in the estimates as a consequence of the coronavirus pandemic; these population estimates have been used to understand both the level of infection in local areas and the success of the vaccine rollout. The accelerated publication schedule has resulted in some changes to the release. Notably, we have prioritised making data available early at the expense of additional primary analysis.

## 6 . Methods used to produce the population estimates data

This section briefly describes the methodologies used to produce small area population estimates (SAPE). These include methods for a standard (non-census) year, methods used for mid-2011 following the 2011 Census and methods for producing the revised mid-2002 to mid-2010 estimates. Full details of all these methods are given in the [Methodology note on production of small area population estimates](#).

## Ratio change methodology - Super Output Areas estimates, mid-2012 onwards

The Lower layer Super Output Area (LSOA) and Middle layer Super Output Area (MSOA) estimates are consistent with the local authority mid-year estimates produced using the cohort component method. Below local authority level, a ratio change method is used to obtain LSOA- and MSOA-level data constrained to local authority estimates.

The main assumption behind this ratio change method is that, for each area, the data have a consistent relationship with the usual resident population over time. Changes in administrative datasets counts, expressed as a change ratio, are applied to a previously estimated population to produce an updated population estimate.

Change ratios are calculated by dividing, for each dataset, the Year 2 count by the Year 1 count for each quinary age or sex group: for example, for female 20- to 24-year-olds, a Year 2 count of 50, divided by a Year 1 count of 40, gives a change ratio of 1.25. The change ratio is then applied to the Year 1 population estimate, to derive a population estimate for Year 2.

Up until the original publication of the mid-2013 estimates, two main administrative datasets were used in the ratio change method: Patient Registers and Child Benefit. Some age groups were covered by a single dataset, while other age groups were covered by both. The datasets used to calculate the LSOA and MSOA change ratios from Year 1 to Year 2 for quinary age groups and sex were as follows:

- 0 to 4, 5 to 9 and 10 to 14 years - Child Benefit and Patient Registers (mid-2012 only)
- 15 to 19 years, to 90 years and over - Patient Registers

From the original mid-2013 estimates and for all years of the revised backseries for mid-2012 to mid-2016, no Child Benefit data was used due to changes in the dataset following the government policy introduced in 2013 to restrict the availability of Child Benefit depending on family income. See the following section on output quality trade-offs for further details.

Before applying these change ratios, some population counts are subtracted (referred to as the special population) comprising UK armed forces, foreign armed forces and prisoners, and added again after these counts are constrained to the Year 2 local authority mid-year estimates minus the special population. These special populations are not covered by the administrative datasets and are therefore treated separately.

For more detail on the ratio change methodology and how it is applied, please see [Methodology note on production of small area population estimates](#).

## Output Area apportionment methodology — estimates for other geographies, mid-2012 onwards

The Output Area (OA) apportionment method provides a mechanism for deriving estimates, which are consistent across all geographical levels. The method is versatile and easy to implement, which helps to overcome the difficulty of producing annual population estimates when boundaries are subject to periodic review and change.

In essence, the method apportions data for LSOAs down to OA level using Patient Register data. An allowance is made for armed forces and prisoners, which are population sub-groups not covered by the Patient Registers. These OA estimates can then be aggregated to a range of higher geographies using a best-fit allocation. This approach is designed to comply with the ONS Geography policy.

A slightly different approach has to be taken for National Parks due to the fact that OA best-fitting is not considered fit for purpose in these areas. However, the estimates are still ultimately based on the OA population estimates.

For more detail on this, please see [Methodology note on production of small area population estimates](#).

## **Mid-2002 to mid-2010 revised estimates (methodology for producing revised OAs, SOAs, wards, Parliamentary constituencies and health geography)**

The mid-2002 to mid-2010 SAPE are revised in line with the results of the 2011 Census. Estimates are consistent with the revised census-based local authority (where appropriate) and national population estimates and reflect the administrative boundaries in place on 30 June 2011. Revised mid-2002 to mid-2010 ward estimates are published on 2012 boundaries.

The estimates are updated to account for population change over the intercensal period, to provide a consistent time series for future estimates based on the 2011 Census.

The census-based mid-2011 OA and LSOA estimates and the mid-2011 rolled forward estimates for each OA and LSOA respectively are used to identify differences that may have arisen over the intercensal period. The observed difference in the intercensal period is equally distributed over time; that is, smoothing the change across the series.

A summary of the method is as follows:

- the existing mid-2002 to mid-2010 OA and LSOA rolled-forward estimates are produced on the new 2011 OA and LSOA boundaries; although no revisions have been applied, 2001 OA rolled-forward estimates will also be produced on the new 2011 OA boundaries to provide a continuous time series from 2001 to 2011
- adjustments are made for corrections for known issues in the mid-2002 to mid-2010 rolled forward estimates
- the mid-2002 OA and LSOA (by single year of age and sex) rolled-forward population estimates are used as the base year
- the difference at OA and LSOA level in mid-2011 is distributed cumulatively across the OA and LSOA backseries respectively by age-sex cohort, for example, 10% of the difference at single year of age and sex is applied to mid-2002, 20% to mid-2003, 30% to mid-2004 and so on

The resulting LSOA estimates are constrained to the previously published revised census-based local authority estimates and the resulting OA estimates are constrained to the revised and constrained LSOA estimates.

MSEA and health geography (primary care organisation (PCO) and clinical commissioning group (CCG)) estimates for mid-2002 to mid-2010 are produced using aggregations of mid-2002 to mid-2010 LSOA estimates. The revised mid-2002 to mid-2010 Output Areas are the building blocks to form the revised mid-2002 to mid-2010 estimates for wards and Parliamentary constituencies.

For more detail on this, please see [Methodology note on production of small area population estimates](#).

## **Mid-2011 estimates (methodology for production in a census year)**

The mid-2011 SAPE are based on the 2011 Census, updated to account for population change during the period between census day (27 March 2011) and the mid-year point (30 June 2011).

The 2011 Census data used as the base are adjusted to include:

- age at the mid-year point (as opposed to census day age given in 2011 Census outputs)
- armed forces at their place of usual residence according to the definition applied for mid-year population estimates (differs slightly from 2011 Census definition)
- prisons in the area in which the majority of the prison buildings fall (again differs slightly from the 2011 Census definition)

For SOAs, the base 2011 Census data are adjusted to account for any births and deaths that have occurred during the period 27 March to 30 June 2011 and the resulting estimates are constrained to the previously published local authority estimates. This constraining process accounts for all other sources of population change including both internal and international migration.

From mid-2011 onwards, estimates for wards and Parliamentary constituencies are produced using an OA best-fit approach in accordance with the ONS geography policy. Mid-2011 OA estimates are derived from LSOA estimates using an apportionment method. This applies the 2011 Census distribution of population between the OAs in each individual LSOA to the mid-2011 LSOA population estimates to break them down to OA level.

For more detail on this, please see [Methodology note on production of small area population estimates](#).

## Statistical disclosure control

The estimates are subject to statistical disclosure control. The general application of statistical disclosure control ensures that information attributable to an individual or organisation is not disclosed in any publication.

The [Code of Practice for Statistics](#) and, specifically, the pillar on trustworthiness, sets out principles for how we protect data from being disclosed. The pillar includes as part of principle T6.4: "Organisations should be transparent and accountable about the procedures used to protect personal data when preparing the statistics and data...appropriate disclosure control methods should be applied before releasing statistics and data...". More information can be found in [National Statistician's Guidance: Confidentiality of Official Statistics](#).

The disclosure control processes applied to the estimates include small adjustments made to selected cells. Adjustments are made in such a way that inference of an underlying count is not possible but that the usefulness of the aggregated estimates is not materially affected.

## 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. Our 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 our website but not produced by us, or referenced on our 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](#)
- [accessibility](#)

In addition to this quality and methodology information, basic quality information relevant to each release is available in the quality and methodology section of the relevant statistical bulletin.

SAPE by age and sex are available for 2001 onwards. The estimates are accompanied by statistical bulletins from mid-2010 onwards. Standard tables can be downloaded free of charge in .xls file formats. More detailed tables, for example, by single year of age, are also published. Metadata describing the limitations and suitable use of the data from these more detailed tables are provided with each individual dataset. Most queries can be answered from the website datasets or supporting methods documents.

To meet user needs for more detailed geographic information the provision of OA estimates allows for the creation of population estimates for most geographic areas using a "best-fit" approach. This gives users the flexibility of being able to obtain population estimates for locally defined or bespoke areas. Information about the different geographies used in England and Wales and "best-fitting" can be found on the [ONS geography](#) website. Any other requests for data that take more than 30 minutes to process will be charged for in accordance with the [ONS Income and Charging Policy](#).

## 7 . Other information

Any additional enquiries regarding the small area population estimates or requests for additional data can be made via email [pop.info@ons.gov.uk](mailto:pop.info@ons.gov.uk) or by telephone on +44 (0)1329 444661.

Advance notice of any forthcoming major changes in methodology will be announced on the ONS website.

### Useful links

Latest mid-year population estimates (national and local authority level):

- [Publications, data tables and methodology information.](#)

Revised mid-year population estimates mid-2012 to mid-2016:

- [Revised population estimates for England and Wales: mid-2012 to mid-2016](#)

Revised mid-year population estimates (mid-2002 to mid-2010):

- [National](#) (England and Wales)
- [Subnational](#) (local authorities)

Population estimates by output area, electoral, health and other geographies:

- [Mid-2020 and links to previous years' data](#)
- [An Overview of ONS Population Statistics.](#)