Quality overview of data sources used in mid-2025 population estimates for England and Wales

This quality overview provides descriptions of the strengths and limitations of all the individual administrative (admin) data sources used directly in producing the mid-2025 population estimates (MYEs) for local authorities (LAs) in England and Wales.

Contact:
Demography team
pop.info@ons.gov.uk
+44 1329 444661

Release date:
29 July 2026

Next release:
To be announced

Table of contents

1. Overview

2. Use of data sources

3. How we assess quality

4. Personal Demographics Service

5. Higher Education Statistics Agency

6. Ministry of Justice prisoners

7. Births and deaths registrations

8. Ministry of Defence, UK Armed Forces

9. United States (US) Air Forces in Europe – UK, Foreign Armed Forces

10. University of Warwick

11. Related links

12. Cite this methodology

1 . Overview

This quality overview report provides descriptions of the strengths and limitations of all the individual administrative (admin) data sources used directly in producing the mid-2025 population estimates (MYEs) for local authorities (LAs) in England and Wales.

The MYEs are accredited official statistics. Detailed information on how the MYEs are derived is provided in our accompanying Quality and methods guide, while this report focuses on the admin sources used as input data for the estimates.

There is a very brief description of how the different data sources feed into the MYEs in Section 2: Use of data sources.

In Section 3: How we assess quality, we set out our standard approach to assessing quality, which is applied to each data source covered in this report.

The main body of this report contains summary information on the quality of each admin data source contributing directly into the MYEs. For this, we have adapted the European Statistical System (ESS) Quality Dimensions of:

- relevance
- accuracy and reliability
- timeliness and punctuality
- accessibility and clarity
- coherence and comparability

Information on the relevance of each data source, and thus why it was selected, is summarised in the introductory paragraphs for each data source. A section on coverage describes how well the data source covers the population of interest and any differences to standard definitions. Summary information on accuracy and completeness, and timeliness is provided in subsequent sections. Any lack of comparability with previous versions of the data is highlighted where relevant.

We welcome feedback on this approach to setting out the strengths and limitations of each source, and whether there are any areas where more detailed information would be useful. Please email us at pop.info@ons.gov.uk.

Other published reports contain similar information for data sources, which indirectly feed into the MYEs. Links to these reports are provided in Section 2: Use of data sources. All the admin data described here have been acquired in line with the Office for National Statistics's (ONS's) Data acquisition policy and Data ethics policy to assist with our research into transforming census and data collection practices. Data supply agreements with the supplier are in place for each source and the data are subject to rigorous controls to ensure that individuals cannot be identified. Further information, including the ONS's privacy statement and data protection policy, can be found on our Data protection web page.

2 . Use of data sources

Detailed descriptions of how the mid-year population estimates (MYEs) are created, and on their quality assurance, appropriate usage, and strengths and limitations, are available in our Mid-year population estimates for England and Wales quality and methods guide.

The use of admin data sources within the method can be summarised as follows.

1. Population changes over the year owing to births and deaths are derived from civil registration data for those events.

2. Population changes over the year owing to migration are estimated using the Office for National Statistics's (ONS's) Long-Term International Migration (LTIM) estimates (in conjunction with data sources used to disaggregate those estimates to local authority level) and estimates of migration within the UK used in the mid-year estimates. Both these sets of estimates themselves rely on administrative data sources, which feed indirectly into the MYEs.

3. Population changes owing to UK Armed Forces personnel are derived using information provided by the Ministry of Defence (MoD).

4. Population changes over the year owing to Foreign Armed Forces personnel are derived using information provided by the US Visiting Forces (USVF).

5. Population changes over the year owing to Ministry of Justice (MoJ) prisoners are derived using information provided by the HM Prisons and Probation Service (HMPPS).

List of data sources

The following data sources are used to produce the Mid-year population estimates:

- Personal Demographic Service (Annual Stocks)
- Personal Demographic Service (Weekly updates)
- Higher Education Statistics Agency (Student Record)
- Death Registrations
- Birth Registrations
- Ministry of Justice (MoJ) Prisoners (Record Level)
- Ministry of Defence (MoD) UKAF Personnel (Aggregated)
- United States (US) Air Forces in Europe – United Kingdom (Aggregated)
- University of Warwick halls of residence data

These sources are covered in this report as they feed directly into the mid-2025 MYEs.

Further information regarding the mid-2025 MYEs can be found in our accompanying Quality and methods guide, in our Long-Term International Migration methodology and in our Subnational long-term international migration methodology: July 2026.

3 . How we assess quality

Every data source used in the mid-year population estimates (MYEs) is subject to continued evaluation. This considers how well the data source meets our requirements in:

- producing estimates
- the governance around the data source and the systems used in processing it
- the quality and sustainability of the data
- the presence of contingency plans if the data are not provided to the expected time or quality

In addition to this general assessment of quality when using data sources in the MYEs, three further steps of evaluation and assurance are conducted for each supply of a selected data source.

1. We hold discussions with the supplier to identify any changes in definitions, collection procedures, quality assurance or similar aspects that might affect the data.
2. We conduct a wide range of standard checks on the structure and validity of the data received, including that data are of the expected format; the degree of missingness for each variable; numbers of invalid or unexpected values (including dates outside an expected range) for each variable; the quality of variables used in linking the data source to other sources; numbers of duplicate records; and evidence of any large clusters of records assigned to single addresses or postcodes.
3. We examine demographic and geographical patterns in the data and compare these with patterns in previous supplies and other sources.

Any anomalies identified in the data are investigated and raised with the data suppliers where needed.

4 . Personal Demographics Service

The NHS Personal Demographics Service (PDS) is the national electronic database that contains demographic data as well as a national unique identifier – the NHS number – for those who have interacted with an NHS service in England, Wales, the Isle of Man, or UK defence medical services, including interactions through GP practices and hospital visits.

We use two types of extract from the PDS in producing the mid-year population estimates (MYEs).

The first type is the "annual stocks" extract, which is a snapshot of the PDS as of 31 July of the year.

The second type is the "weekly update" files, which record changes of address. These files are used in estimating internal migration moves, which feed into the MYEs.

The PDS stocks and PDS updates are used in the calculation of internal migration flows, and this process is explained in our accompanying Quality and methods guide.

The PDS stock extract is also used in the production of Long-Term International Migration estimates to distribute international migrants, both EU and non-EU, under the age of 16 years, to local authority areas. More on administrative data used in the production of international migration estimates can be found in our Long-term international migration: quality assuring administrative data methodology and in our Subnational long-term international migration methodology: July 2026.

Coverage

The PDS has a very high coverage of the usually resident population at all ages at the national level (England and Wales). Records are created for newborns or when a patient contacts an NHS service, primarily by registering with a General Practitioner (GP), but also through accessing Accident and Emergency (A&E) or attending hospital. Further PDS records are created from information collected and provided by the Home Office for visitors and migrants who have registered under the European Union Settlement Scheme, who have paid the Immigration Health Surcharge, or are part of a vulnerable cohort (such as asylum seekers and victims of trafficking).

People having private health care and irregular migrants are less likely to be included, as seen in our Administrative data used in Census 2021, England and Wales methodology. It is believed that only a small minority of the total population is not known to the NHS and therefore not included in the PDS data. The data delivered contain approximately 85 million patient records of which over 60 million are for people currently registered with a GP.

We are also aware of a level of overcoverage in the PDS data. The data contain information on the reasons individuals can be removed from the database such as if they relocated within the UK, embarked abroad or died. We use this information to flag around 15 million records, unlikely to relate to a usual resident. However, not everyone is immediately removed from the system, for example, when usual residents emigrate without informing the NHS.

We are working with NHS England, and investigating further methods, to better understand and address overcoverage in the data.

Accuracy and completeness

The data have a high level of completeness for the population covered, with low levels of missingness for variables used in the derivation of the MYEs.

Interactions with PDS do not generally occur on an annual basis for most of the usually resident population. Records are only updated when an individual updates their details at a GP or through other interactions with health services via forms completed online or on paper under the direction of departmental staff. Different collection methods, particularly paper forms, can lead to an increase in entry error, therefore reducing the accuracy of the data.

However, NHS England has developed a new GP registration service that provides all GP practices with an integrated online option for patients. This allows online updates to contact details to improve the accuracy of the PDS and reduce third-party input errors.

Overall, the accuracy is usually high for age and sex, but address information might be affected by delays in people updating their records. Comparisons of the PDS with Census 2021 found that the PDS tends to be more accurate for younger ages (schoolchildren and particularly those aged under 1 year old) and older ages (those aged 55 years and over), whereas student ages and young men aged in their 20s and 30s are less likely to interact with public services on a regular basis, as was reported in previous findings from ONS's methodological work for producing census-based 2021 mid-year estimates (MYE).

Similarly, findings from our linkage between Census 2021 and the PDS carried out in 2023, suggest that the accuracy of the PDS data on address varies across LAs between England and Wales. In areas which have higher population churn, for example urban areas and those with more higher education institutions, the PDS tends to be less accurate. This lower level of accuracy can lead PDS to show some overcoverage in more urban LAs where there is a higher rate of people moving in and out of the LA. For example, those visiting an LA may temporarily interact with an NHS service and then move out of the area but remain on the PDS (for example, short-term migrants).

Timeliness

The annual stocks file relates to the PDS data as they were on 31 July of that year. This reference date accounts for the typical one-month delay in GP registration following a move. This means that the MYEs' reference date of 30 June will be less affected by a lag in updating records than if the same reference date were used for the PDS extract. The 2025 PDS stocks extract was delivered to the Office for National Statistics (ONS) in August 2025 and thus provides very timely information that aligns well with the MYEs mid-year point.

Weekly update files are also received in a timely manner, supporting the MYEs production schedule.

5 . Higher Education Statistics Agency

The Higher Education Statistics Agency (HESA) is part of Jisc, the UK's digital, data and technology agency for tertiary education, that collects and disseminates data on higher education (HE) across the UK. Jisc is the designated data body for England under the Higher Education and Research Act 2017. The HESA Student Record data are submitted by higher education providers for all students attending a course that leads to the award of a higher education qualification or higher education-level credit. This includes those awarded qualifications by the provider or another awarding body, including international students.

The data are primarily used to supplement Personal Demographics Service (PDS) data when producing internal migration estimates, since they provide better quality data on residential location of a large group of young healthy adults, particularly males, that tend to be slow in updating their details in other data sources.

Our Long-term international migration: quality assuring administrative data methodology provides more information and assurance about the quality of the HESA Student Record data used to produce migration estimates of EU nationals.

HESA data are also used in conjunction with Warwick University Halls of residence data as part of the Coventry /Warwick adjustment.

Coverage

The HESA Student Record includes data from HE providers with a statutory duty to report to funding and regulatory bodies across the UK. Although some privately funded institutions are not required to provide these data, the HESA data provide strong coverage of students in HE, including international students residing in England and Wales.

Those studying overseas or in the UK for fewer than eight consecutive weeks are not included in the HESA Student Record. An exception applies to a small number of distance learners, such as Crown servants overseas or members of the British armed forces, whose tuition is funded by a UK government department, or who are eligible for a tuition fee loan.

The HESA Student Record includes anyone in active study at least two weeks after their start date (or anniversary of it) during the reporting period (1 August to 31 July). This may result in some overcoverage, as students who drop out after the start of the term remain in the data, even if they no longer reside at their term-time address.

Accuracy and completeness

The data have a high level of completeness for the population covered, with low levels of missingness for variables used in the derivation of the mid-year estimates (MYEs).

HESA collects data from a large number of HE providers, who are required to report to their funding and regulatory bodies. To support this, HESA provides guidance and resources to ensure consistency. All submissions undergo automated validation checks, and any that fail are returned to providers for correction. Therefore, HESA Student Record data are considered a high-quality source.

However, there are some limitations when using MYEs. Student address information may become less accurate over time. Permanent home postcodes are collected at the start of study, while term-time postcodes are updated annually with each course session. Practices vary across and even within HE providers.

Students can move to a new term-time address shortly after registering their details with their HE provider, but these updates may not be reflected in the HESA data. This means that the address recorded within the HESA Student Record may not reflect the student's address on 30 June, the MYEs reference point.

The impact of this is not thought to be substantial because most in-term student moves are usually short-distance and remain within the same local authority. Since MYEs are produced at local authority level, such moves are unlikely to affect the statistics. Assuming these moves are not systematically patterned, any random errors in these tend to cancel out. Additionally, when HESA Student Record data are used for internal migration estimates, we supplement them with more recent address information from other data sources like patient registrations.

Timeliness

The extract for the academic year ending 2025, used in the 2025 MYEs, was delivered to the Office for National Statistics (ONS) in February 2026. The bulk of data collection is done at the start of the academic year by the HE providers with a lag of around 18 months from this point to the data being delivered to the ONS. However, the HE providers update records throughout the year ahead of the final submission to HESA, including where students alert them of a change (for example, an address change). The period of collection covers 1 August 2024 to 31 July 2025, with the reference date for the MYEs being 30 June 2025. We are working with Jisc to improve the timeliness of the data.

6 . Ministry of Justice prisoners

These data are supplied to the Office for National Statistics (ONS) by HM Prison and Probation Service (HMPPS), an executive agency sponsored by the Ministry of Justice (MoJ). The MoJ processes person-level prisoner data for law enforcement purposes including processing for execution of criminal offences as provided under Part 3 (Law Enforcement Directive) of the Data Protection Act 2018 (DPA).

The data cover all prison establishments in England and Wales, which are required to record prisoner details on the Prison National Offender Management Information System (p-NOMIS). They include personal information, such as prisoners' name, date of birth, age and sex, that allow them to be linked to other admin data sources.

The data provide valuable information on the current residence of this population group and are the basis of the special population adjustment for population estimates.

Coverage

The data provide excellent coverage of individuals held in custody, including all prisoners aged 15 to 21 years in public and private young offender institutions (YOIs), and adult establishments in England and Wales, under the following legal status groupings:

1. Sentenced
2. Recall
3. Remand
4. Indeterminate Sentence
5. Convicted Unsentenced
6. Immigration Detainee
7. Unknown
8. Civil Prisoner
9. Other

The data do not cover individuals held at immigration removal centres as these are non-criminal detainees (unlike immigration detainees housed in prisons).

Accuracy and completeness

The data have a high level of completeness for the population covered, with low levels of missingness for variables used in the derivation of the mid-year population estimates (MYEs).

Data on prisoners are updated in real time by the various institutions involved in their collection and kept up to date, meaning that the data are accurate and do not suffer from significant lag effects.

Confidence in data accuracy and reliability is also believed to be high because of the status of its data subjects, as they are a population from which administrative information can be efficiently collected.

Timeliness

The annual extract from p-NOMIS reflects the data as of 30 June each year, aligning with the MYEs' mid-year reference point.

The 2025 p-NOMIS extract, used in the 2025 MYEs, was delivered to the ONS in August 2025. This represents a short lag between the reference date and delivery, providing timely data for inclusion in the MYEs.

7 . Births and deaths registrations

Births and deaths registrations data are sourced from the Local Registration Service in partnership with the General Register Office (GRO), which records all births and deaths in England and Wales. These data are used in mid-year populations estimates (MYEs) to estimate population change because of births and deaths. More information on the quality of the births and deaths registrations, can be found in our Births quality and methodology information report and in our Mortality Statistics in England and Wales quality and methodology information report.

Coverage

The Civil Registration System provides very high coverage of births and deaths occurring in England and Wales. Elements of undercoverage resulting from births and deaths happening to usual residents while abroad are assumed to be offset by corresponding overcoverage from those events happening within England and Wales to people who are not usual residents.

The data used in the mid-2025 MYEs relate to births and deaths occurring in the 12 months to end of June 2025 and registered at any time up to April 2026. This will effectively cover all births. A small adjustment is made to reflect the estimated number of deaths occurring in the MYE reference period but not appearing in the registration data. More information on late registration of deaths is provided in our Impact of registration delays on mortality statistics dataset.

Accuracy and completeness

The data are considered highly accurate and complete because of the legal framework and robust verification processes in place. Under the Births and Deaths Registration Act 1953, there is a legal requirement in England and Wales for parents to register a birth within 42 days, and for deaths to be registered within five days unless the death is referred to a coroner or there are exceptional circumstances, for example, stillbirths. This legal obligation ensures a high level of compliance and data reliability.

While birth notifications - initial records typically completed by a midwife or doctor shortly after birth - are timelier, registrations are more complete and of higher quality. For birth registrations, the informant, usually the mother or both parents, must verify the information provided to the registrar. The data are entered into the Registration Online (RON) system, which includes built-in validation checks to reduce entry errors. We carried out additional quality assurance upon receipt of the data. Supplying false information is a criminal offence, hence birth registrations accuracy is generally believed to be high. The accuracy of sex and date of birth is particularly high, while names and addresses are generally reliable, though minor spelling or formatting inconsistencies may occur.

For death registrations, the informant, typically a close relative, hospital official, or coroner's officer, must verify the information before submission. Sex is recorded based on the Medical Certificate of Cause of Death (MCCD) and is considered highly reliable. Names and addresses are generally accurate, though address information may be affected by recent moves or temporary residence at the time of death.

Timeliness

Death registrations data are supplied to the Office for National Statistics (ONS) daily when a registration is made through the RON system. As described in the Coverage section, the registration data used in the MYEs align very well with the reference period of the estimates, with a small adjustment made for very late registration of deaths.

8 . Ministry of Defence, UK Armed Forces

The Ministry of Defence (MoD) supplies the Office for National Statistics (ONS) with data on military personnel based both in the UK and in the rest of the world.

Home Armed Forces (HAF) in UK data are supplied by the Defence Statistics (Tri-Service), and HAF in Germany data are supplied by Strategic Command.

The Tri-Service is a department within MoD responsible for Armed Forces personnel Official Statistics.

The HAF Germany data provides record-level data on the number of UK armed forces personnel and their dependants based in Germany.

Coverage

Home Armed Forces (HAF) data provide full coverage of all armed forces personnel. The HAF Germany data provide full coverage of armed forces personnel and their dependants in Germany.

Accuracy and completeness

The data supplied are complete and accurate as the data are derived from administrative records held by the MoD for operational and personnel management purposes. Data are provided using the local authority location of the base that military personnel are assigned (their workplace) and we use Census 2021 to redistribute military personnel from their place of work to their place of residence.

Timeliness

Home Armed Forces in the UK

Data are delivered on a quarterly basis in approximately February, May, August, and November, based on extracts from MoD's systems from January, April, July, and October. The extract from July each year is used in the production of the mid-year population estimates (MYEs); extracts from January, April and October are used to understand the representativeness of data for July.

Home Armed Forces in Germany

Data are delivered on a quarterly basis in approximately January, April, July, and October, based on extracts from MoD's systems from March, June, September, and December. The extract from June each year is used in the production of the MYEs; extracts from December, March and September are used to understand the representativeness of data for June.

9 . United States (US) Air Forces in Europe – UK, Foreign Armed Forces

The United States (US) Air Forces in Europe – UK (USAFE-UK) supplies the Office for National Statistics (ONS) with annual aggregated counts of US military personnel and their dependants based in the UK. US Armed Forces are treated as a special population – termed the Foreign Armed Forces (FAF) – in the calculation of the mid-year estimates, as the movements of military personnel are not captured by the usual data sources used to estimate international and internal migration.

Coverage

These data are supplied aggregated by age, sex, base-location and personnel category.

Accuracy and completeness

These data are regularly subject to missingness in the form of missing location, missing age and missing sex. In these situations, we contact the supplier to understand the reasons for missingness and to learn the appropriate means to mitigate the issue. Often this involves imputing missing data using the observed patterns in the data we receive or by imputing missing data using the observed patterns in previous years' supplies. Data are provided using the base that military personnel and their dependants are assigned to (their workplace) and we use Census 2021 to redistribute them from their place of work to their place of residence.

Timeliness

The extract for the mid-year 2025, used in the 2025 mid-year population estimates (MYEs), was delivered to the Office for National Statistics (ONS) in July 2025.

10 . University of Warwick

These data are supplied to the Office for National Statistics (ONS) by the University of Warwick. We use the University of Warwick halls of residence data to make an adjustment between the population estimates for Coventry and Warwick. University of Warwick halls of residence all have the same postcode, even though the halls are partly located in Warwick and partly in Coventry. These data are needed to make an adjustment to better reflect the place of usual residence of students.

Coverage

There is high coverage of students living in university-owned halls that are included in the accommodation allocation system, as the source is derived from university records.

Accuracy and completeness

The data provide complete information on the number of students by hall of residence on the University of Warwick campus. However, there are some acknowledged limitations when applying the data to our uses: the age and sex information for each hall of residence is not available, but the age and sex distribution of the students can be imputed by using data from the Higher Education Statistics Agency (HESA) for students living in halls of residence at the university.

Some campus bed spaces can be used by staff rather than students; a total number of staff is provided with the data, which are distributed among the halls and removed to avoid double-counting.

Some bed spaces are occupied by short-term residents, such as those attending courses of less than one year. Therefore, the bed-space figures provided have to be reduced by 20% to estimate bed spaces occupied by students resident for 12 months or more (and therefore meeting the definition of "usually resident").

Timeliness

The extract for the academic year ending 2025, used in the 2025 mid-year populations estimates (MYEs), was delivered to the Office for National Statistics (ONS) in February 2026.

11 . Related links

Mid-year population estimates for England and Wales quality and methods guide
Methodology | Released 29 July 2026
How we produced the statistics, their strengths and limitations and further quality information.

Local authority quality of population estimates for England and Wales
Supporting methodology article | Released 29 July 2026
Explores how uncertainty, as a measure of quality, can be introduced into the population estimates process, and the impact of this on local authorities.

Effect of the coronavirus pandemic on data and quality of population estimates for England and Wales
Supporting methodology article | Released 29 July 2026
An assessment of how the coronavirus (COVID-19) pandemic affected the production and quality of population estimates for England and Wales.

Improving population estimates for England and Wales
Supporting methodology article | Released 29 July 2026
Outline of our continuous improvement plan for mid-year population estimatesfor England and Wales.

12 . Cite this methodology

Office for National Statistics (ONS), released 29 July 2026, ONS website, methodology article, Quality overview of data sources used in mid-2025 population estimates for England and Wales