Developing population estimates by ethnicity for England and Wales

Research into methods for producing annual population estimates by ethnic group and results from testing against Census 2021. This is a research publication and does not present official statistics.

Contact:
Population and Migration Research and Development team
pop.info@ons.gov.uk
+44 1329 444661

Release date:
6 November 2026

Next release:
To be announced

Table of contents

1. Overview

2. Data sources

3. National and regional methods

4. National and regional estimates

5. Local authority methods

6. Local authority estimates

7. Future developments

8. Provide feedback

9. Definitions

10. Data on developing population estimates by ethnicity

11. Related links

12. Cite this page

1 . Overview

The latest official population estimates by ethnic group for England and Wales are from Census 2021. There has been a longstanding need for more up-to-date and regular estimates to better understand how the ethnic composition of the population changes over time. This is important to a wide range of users, including national and local government, to help inform policy making and the planning and provision of services.

The Office for National Statistics (ONS) is researching methods for producing robust population estimates by ethnic group between censuses. We have developed a ratio change method for national and regional estimates, which produced estimates at the five-category ethnic group level that generally showed close agreement with Census 2021.

For local authority estimates, we have continued testing two estimation models, the Structure Preserving Estimator (SPREE) and the Generalised Structure Preserving Estimator (GSPREE). Our testing shows that using the national estimates produced by ratio change as inputs for local authority estimates in GSPREE produces the closest agreement with Census 2021 and improves upon previous testing of this method.

Overall, across national, regional and local authority geographies, the population estimates by ethnic group produced still suffer from the main underlying issues in the individual data sources. The White ethnic group tends to be overestimated while ethnic minorities tend to be underrepresented, with the Other ethnic group being the most difficult to estimate accurately. However, estimates produced using the methods tested largely improve upon the degree of bias present in the individual sources.

The next stage of the research will include engagement with users to understand whether the quality of the estimates meets their needs and to gather requirements to inform further methodological development. This work represents proof-of-concept research for 2021, which allows estimates produced using these methods to be compared with Census 2021. The results are not official statistics and are intended to support method evaluation and development. They should not be used for policy or decision-making purposes.

These are not official statistics and should not be used for policy or decision making. Census 2021 provides accredited official population estimates by ethnicity.

2 . Data sources

The admin-based ethnicity dataset

Administrative data are collected through the day-to-day delivery of public services and government functions rather than for statistical purposes. As part of our aim to provide the best possible insights on population and society using a range of data sources, we have been developing the Admin-Based Ethnicity Dataset (ABED) since 2020.

The ABED is a de-identified record-level dataset that links ethnicity data from 11 administrative data sources, including health and education records, and Census 2011. Sources are linked through the Statistical Population Dataset (SPD 5.1).

A set of rules are applied to resolve instances where multiple ethnicities are recorded for the same individual. ABED captures information for 18 ethnic groups, however, our research has so far used the higher-level five ethnic group categories (Asian, Black, Mixed, Other and White), as the data are of sufficient quality at this level to test the methods explored in this research.

For 2021, ABED Version 5 contains records for 57.7 million people in England and Wales in 2021, compared with 59.7 million people recorded in Census 2021. This means ABED includes approximately 2.0 million fewer people than Census 2021. Of the 57.7 million people included in ABED, ethnicity information is available for around 48.0 million people (83.2%). Compared with Census 2021, ethnicity information is therefore unavailable for around 11.8 million people, meaning ABED provides ethnicity coverage for around 80% of the population.

For individuals in ABED with a recorded ethnicity, ABED captures some of the population distributions accurately. For example, there is close alignment between ABED and Census 2021 for those identifying as White: British. However, representation of ethnic groups differs, meaning the people in ABED with recorded ethnicity are not fully representative of the population. For example, Other ethnic group: Arab is notably underrepresented in ABED.

Quality issues within ABED largely stem from high levels of "Not provided", "Unknown" or "Unresolved" ethnicity classifications and challenges associated with the complexity of how ethnicity data are captured across administrative sources. These factors mean that ABED cannot be used on its own to produce representative population estimates by ethnic group. As a result, we are researching estimation approaches that combine ABED with other data sources to produce annual population estimates by ethnic group for England and Wales.

ABED has been developed through an ongoing programme of research since 2020. Initial work focused on assessing the feasibility of producing ethnicity statistics from linked administrative data for England. Since then, the dataset has been expanded to include England and Wales, and important developments have included the addition of education data sources, such as the Individualised Learner Record, the Welsh School Census and the Lifelong Learning Wales Record and updating the population base to SPD Version 5.1.

As part of future developments to ABED, we are adding Census 2021 as an additional data source, alongside Census 2011, and assessing alternative ethnicity selection rules to improve the quality and coverage of ethnicity information within the dataset. These developments have not been applied to the version of ABED we are currently using in our research.

Census

Census data are the most comprehensive source of information on ethnic group in England and Wales. Conducted every 10 years, the census provides a detailed snapshot of the resident population living in households and communal establishments. Census 2021 provides the latest official population estimates by ethnic groups and serves as an important benchmark for understanding the ethnic composition of the population.

However, census data alone cannot provide up-to-date population estimates by ethnicity beyond the census year. Population change is influenced by births, deaths and migration, and these changes do not affect all ethnic groups equally. As time passes following a census, the ethnic group distributions seen in a census become less representative of the current population.

We previously produced Population estimates by ethnic group (PEEGs) for each year from 2001 to 2009, which combined census-based populations with information on births, deaths and migration. These estimates were discontinued following the publication of the mid-2009 estimates in May 2011 because of concerns about their accuracy. Evaluation against Census 2011 showed that they produced implausible patterns of change over time and were inconsistent with other available sources, including estimates from the Annual Population Survey (APS). This reflected the difficulty of assuming that ethnic group population distributions remain stable throughout the intercensal period.

Producing reliable annual ethnicity estimates requires methods that go beyond carrying forward census distributions or relying on assumptions about how populations change over time. More timely data sources are needed to better reflect changes in the ethnic composition of the population between censuses.

Annual Population Survey

The Annual Population Survey (APS) is a continuous household survey covering the UK. It provides information on a range of social, demographic and labour market characteristics, including ethnic group, employment, housing, health, religion and educational attainment.

The APS combines data from the Labour Force Survey (LFS) with additional sample boosts, making it one of the largest household surveys in the UK. It should be noted that the APS does not cover most people living in communal establishments.

The APS provides more timely information on ethnic group than the census. It has previously been used to produce three-year population estimates by ethnic group for 2014 to 2016, providing information on changes between censuses.

However, the quality of APS estimates has been affected by falling survey response and reduced sample sizes. These challenges intensified during and after the coronavirus (COVID-19) pandemic. Smaller samples reduce the precision of estimates and increase uncertainty, particularly for small areas, small population groups and detailed breakdowns. More information is available in our Annual Population Survey quality update: September 2026.

In this research, APS data contribute to the production of national and regional estimates for 2021 using the ratio change method. APS estimates are used to measure changes in ethnic group distributions between 2011 and 2016, before ABED data became available. For local authority estimates in 2021, APS ethnicity counts are used alongside ABED within GSPREE, with the national ratio change estimates used as benchmarks.

We are transitioning from the LFS and APS to the Transformed Labour Force Survey (TLFS), which is intended to replace APS-based outputs when evidence shows that it can meet user needs and quality requirements. As this develops, we will assess the quality of the TLFS and its potential use in producing population estimates by ethnic group.

3 . National and regional methods

Ratio change

We explored the ratio change method for producing ethnicity estimates and found it performed best at national and regional levels. The method has not previously been applied to ethnicity estimation; it is already used in the production of official small-area population estimates for Middle Layer Super Output Areas (MSOAs) and Lower Layer Super Output Areas (LSOAs).

The method starts with census-based population estimates and updates them over time using changes seen in administrative data. Ratio change is a longitudinal approach. Estimates for one year are based on estimates from the previous year and are updated using changes seen in data between consecutive years.

Our previous applications of the method have typically used a single administrative data source. In this research, we apply the method using the Admin-Based Ethnicity Dataset (ABED), a linked administrative dataset that combines information from multiple sources.

Applying the ratio change method to ethnicity estimation using ABED introduces additional considerations. As with many linked administrative datasets, coverage and data quality can vary across population groups. The ratio change method assumes that changes seen in the administrative data reflect real changes in the ethnic composition of the population over time.

We produced 2021 ethnicity estimates using Census 2011 estimates as the baseline population. As ABED is only available from 2016 onwards, we used Annual Population Survey (APS) estimates to calculate change ratios between 2011 and 2016, and ABED to calculate change ratios between 2016 and 2021.

Estimates produced using only the APS to calculate the year-on-year change for the whole decade resulted in greater divergence from Census 2021. As a result, the approach currently relies on both survey and administrative data. As additional years of ABED data become available, it will be possible to assess the method using a longer administrative data series, reducing reliance on APS estimates.

In the longer term, once a decade of ABED data is available between censuses, the approach could be applied using administrative data across the full estimation period. Census 2031 will provide an opportunity to assess the quality of estimates produced using this extended time series.

Method step by step

At national level, for each ethnic group, a change ratio is calculated by dividing the administrative or survey count in year *t* by the corresponding count in year *t-1*.

For example, if the White ethnic group population recorded in ABED increases from 1,000,000 in one year to 1,050,000 in the next, the resulting change ratio is 1.05.

The ratio is then applied to the benchmark population estimate from the previous year:

$$\text{Estimate}_{t} = \text{Estimate}_{t-1} \times \frac{\text{Admin Data}_{t}}{\text{Admin Data}_{t-1}}$$

where:

- Estimate is the ethnic group estimate for year *t*
- Estimate is the ethnic group estimate for the previous year
- Admin Data is the administrative count for year *t*
- Admin Data is the administrative count for year *t-1*

The estimates in year *t* are constrained to the mid-year estimate (MYE) population total for year *t*, using the MYE series rolled forward from Census 2011. This research uses mid-year population estimates rolled forward from the 2011 Census so the ethnicity estimates are produced entirely independently from Census 2021, enabling us to make valid comparisons between methods.

For regional estimates, the ratio change method is applied separately within each region. The resulting estimate for year *t* is then constrained to the corresponding regional mid-year estimate (MYE) population total for year *t* (rolled forward from Census 2011).

In summary

- Census 2011 ethnic group estimates are used as the baseline population.
- Administrative or survey counts are obtained for consecutive years by ethnic group; for this analysis we used APS estimates for years 2011 to 2016, and ABED estimates for years 2016 to 2021.
- Change ratios are calculated by dividing the count in year *t* by the count in year *t-1*.
- The ratios are applied to the previous year's ethnic group estimates to produce updated estimates.
- The resulting estimates are constrained to the official mid-year estimate (MYE) population totals for the corresponding year; for this analysis we use the MYEs rolled forward from mid-year 2011.
- The process is repeated sequentially for each subsequent year.

Strengths and limitations

An important strength of the ratio change approach is that it uses census data as a starting point. These data provide the most comprehensive measure of the population by ethnic group available for England and Wales. Rather than relying directly on ethnicity distributions seen in administrative or survey data, the method uses changes seen in those sources to indicate how the population may have changed over time. As a result, the estimates are less affected by potential biases in the administrative and survey data, provided these biases remain broadly stable over time.

However, the quality of ratio change estimates is expected to decline as they move further away from the census base year. As 2021 is 10 years after the 2011 Census, the estimates presented in this research are produced at the furthest point from the census base year. As a result, they are likely to be subject to greater uncertainty than estimates produced earlier in the decade.

This is a common challenge for methods that use a census-based starting point and produce estimates over time. Previous applications of the ratio change method, including research to produce population estimates for MSOAs and LSOAs, have been subject to this same limitation. The impact of moving further from the census base year may be particularly important for smaller ethnic groups and for lower levels of geography, where estimates are typically more sensitive to data quality issues and changes in underlying population patterns.

4 . National and regional estimates

National estimates

We produced national-level estimates for the five high-level ethnic groups in England and Wales for 2021. We compared these estimates with Census 2021, the APS and ABED to assess how closely they reflected the ethnic composition of the population.

Table 1: Ratio change method produces national estimates generally closer to Census 2021 than APS and ABED for England and Wales

Percentage of people in each of the five high-level ethnic groups by data source, England and Wales, 2021

Ethnic Group	Census 2021	Ratio Change	Admin-based Ethnicity Dataset (ABED)	Annual Population Survey (APS)
Asian	9.3	9.1	8.6	8.3
Black	4	4	4	3.6
Mixed	2.9	3.1	2.8	2.3
Other	2.1	1.5	1.1	1.3
White	81.7	82.2	83.4	84.4

Source: Office for National Statistics

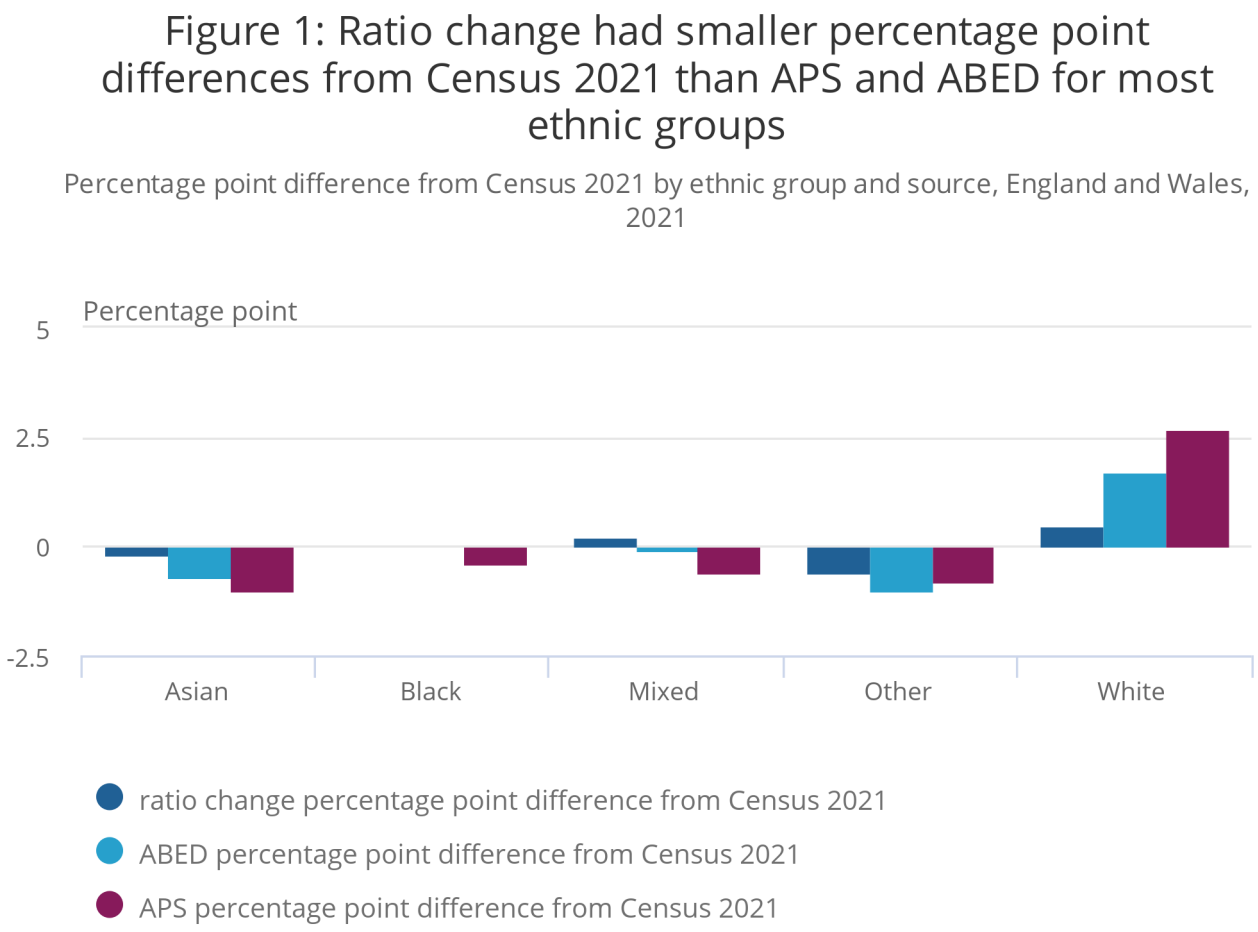
Notes

1. ABED is Admin-Based Ethnicity Dataset. APS is Annual Population Survey.
2. APS percentages are calculated using APS data from January to December 2021.
3. APS estimates exclude most people living in communal establishments, whereas Census 2021 and ABED include people living in communal establishments. This should be considered when interpreting differences between sources.
4. Percentages are rounded to one decimal place.
5. Percentages may not sum to 100% because of rounding.

Table 1 shows that the ratio change estimates were generally closer to Census 2021 than both APS and ABED across the five high-level ethnic groups.

Figure 1: Ratio change had smaller percentage point differences from Census 2021 than APS and ABED for most ethnic groups

Percentage point difference from Census 2021 by ethnic group and source, England and Wales, 2021



Source: Office for National Statistics

Notes:

1. ABED is Admin-Based Ethnicity Dataset, APS is Annual Population Survey.

2. Percentage point differences are shown to one decimal place.

Figure 1 shows that ratio change, ABED and APS produced higher estimates for the White ethnic group and lower estimates for the Asian and Other ethnic groups than Census 2021. However, the differences from Census 2021 were smaller for the ratio change estimates.

For the Black ethnic group, both ratio change and ABED produced estimates that matched the Census 2021 percentage, to one decimal place. For the Mixed ethnic group, the ratio change estimate was slightly higher than the Census 2021 percentage, while APS and ABED both produced lower estimates. This level of agreement is especially notable because the ratio change estimates represent the end of the 10-year period since Census 2011, where quality is weakest and we would expect differences to be greatest.

Regional estimates

We produced 2021 estimates of the distributions for the five high-level ethnic groups in England and Wales for the nine English regions and Wales. Further breakdowns for Wales are included in our local authority analysis.

Overall, the ratio change estimates were generally close to Census 2021, although the size of the differences and whether estimates were higher or lower than Census 2021 varied by ethnic group and region. The ratio change method generally performed better at regional level than estimates produced directly from APS and ABED when compared with Census 2021.

Figure 2: The ratio change method produces estimates at regional level that are broadly similar to Census 2021

Percentage point difference of ratio change estimates compared with Census 2021 for each region, England and Wales

Downloac

Figure 2 shows the difference between proportions estimated by ratio change and proportions from Census 2021 at a regional level. Most differences across regions are less than 1 percentage point.

For the Asian ethnic group, differences between ratio change estimates and Census 2021 showed a tendency for underestimation. Differences ranged from negative 0.9 percentage points in London to positive 0.5 percentage points in the West Midlands.

The Black ethnic group showed the closest agreement with Census 2021, with differences ranging from negative 0.3 percentage points to positive 0.4 percentage points across regions, with most regions showing underestimation.

For the Mixed ethnic group, most differences were small, although estimates were 1.1 percentage points higher than Census 2021 in London and 0.6 percentage points higher in the East Midlands. Compared with the other ethnic groups, estimates for the Mixed ethnic group showed the most even split between underestimation and overestimation, with four regions showing underestimation and six showing overestimation.

The Other ethnic group was underestimated in most regions, particularly in London (negative 1.4 percentage points), the West Midlands (negative 0.8 percentage points), and the East Midlands and South East (both negative 0.6 percentage points).

In contrast, the White ethnic group tended to be overestimated, with the largest differences occurring in London (positive 1.5 percentage points) and the South East (positive 1.4 percentage points).

Based on differences in the distributions by ethnic group, the North East generally showed the closest agreement with Census 2021. The largest percentage point differences were in London, however, the estimates still broadly reflected London's ethnic group distribution. In Wales, estimates were also close to Census 2021 (less than 0.6 percentage point difference) across all ethnic groups.

It is important to note that this analysis focuses on comparing the differences in the population distribution by ethnic group between the ratio change method and Census 2021. The overall size of the population, and the share of the population in each ethnic group varies across regions. As a consequence, the larger differences seen in Figure 2 for London in part reflect the fact that the ethnic composition of London is most different to the national picture compared with other regions.

These findings should also be considered in the context of 2021, 10 years after the 2011 Census and therefore the point furthest from the census base year used in this research. As a result, the ratio change regional estimates discussed in this section are likely to be of lower quality than estimates produced earlier in the decade.

5 . Local authority methods

SPREE and GSPREE

We used the Structure Preserving Estimator (SPREE) and Generalised Structure Preserving Estimator (GSPREE) to produce local authority ethnicity estimates. We did not apply ratio change directly at local authority level in this research.

The smaller sample sizes in the Annual Population Survey (APS) can result in greater volatility and, in some cases, zero values for particular ethnic group and local authority combinations, making it difficult to measure change reliably over time.

SPREE is a small area estimation method that uses two data inputs:

1. a detailed dataset measuring the characteristic of interest for the target population but that may be out of date, incomplete or contain non-equivalent definitions (known as the proxy)

2. a dataset containing trusted totals for the population and characteristic of interest

SPREE then adjusts the proxy dataset so that the estimates match the trusted totals while preserving the structure of the proxy dataset as much as possible.

GSPREE is an extension of SPREE that incorporates an additional survey data source. The survey dataset is also a detailed dataset of the target population and characteristic of interest, and is used to apply an additional adjustment to the proxy dataset before these estimates are constrained to match the trusted totals. We published a research article outlining the SPREE and GSPREE methods in 2017.

When applied at regional level, SPREE and GSPREE models considered so far produced poorer results.

Method step by step

Both methods used:

- the Admin-Based Ethnicity Dataset (ABED) as the proxy dataset
- 2021 mid-year estimates rolled forward from 2011 to provide local authority population totals
- national ethnicity totals produced using the ratio change method

GSPREE additionally used unweighted Annual Population Survey (APS) ethnicity counts for each local authority. SPREE uses iterative proportional fitting (IPF) to adjust the ABED distributions so that they are consistent with both the local authority population totals and national ethnicity totals. In practice this means the population totals (the rows of the proxy dataset) and the ethnicity counts (the columns of the proxy dataset) are repeatedly adjusted in turn until both sets of totals match the trusted totals.

GSPREE first fits a log-linear model to estimate the relationship between the survey and proxy data. This relationship is used to adjust the ABED distributions towards the survey data before constraining the resulting estimates to the trusted population totals using IPF. This allows GSPREE to account for proportional differences in the underlying distribution of the population by local authorities and ethnic group while retaining the detailed information available in the proxy dataset.

The output from both methods is a set of local authority by ethnic group estimates that are consistent with the provided population and ethnicity totals.

SPREE: strengths and limitations

An important strength of SPREE is that it preserves the detailed ethnicity-by-geography patterns seen in the proxy dataset while ensuring consistency with provided population and ethnicity totals.

However, SPREE assumes that the relationship between ethnicity and geography seen in ABED is representative of the wider population. As with many administrative datasets, coverage and data quality can vary across population groups. If some ethnic groups are over- or underrepresented in ABED, these patterns may be reflected in the resulting estimates. The quality of the estimates therefore depends on both the quality of the proxy dataset and the population totals used.

SPREE was included in this research because of concerns over the quality of the survey data available and to provide a useful comparison with the more complex GSPREE method.

GSPREE: strengths and limitations

An important additional strength of GSPREE is that it combines the strengths of administrative and survey data. By incorporating survey data, it reduces reliance on the assumption that the proxy dataset perfectly represents the target population. However, the extent of this benefit depends on how well the survey data represent the distribution of the target population.

The quality of the estimates depends on both the proxy dataset and the survey data used. While survey data can help adjust for biases in the proxy dataset, sample sizes for some ethnic groups and local authorities can be small, increasing uncertainty in the survey-based adjustments. As with SPREE, limitations in the administrative data may also influence the resulting estimates.

A further strength of both SPREE and GSPREE is that they produce estimates separately for each year using the data available for that year. This means the quality of the estimates is not expected to be affected as much by how far they are from a census year. In contrast, methods that estimate change over time from a census baseline, such as ratio change, are expected to reduce in quality as the time since the census increases.

6 . Local authority estimates

Local authority estimates at five-category ethnic group level

We produced local authority estimates for the five high-level ethnic groups in England and Wales for 2021 using SPREE and GSPREE. Both methods used national ethnicity benchmarks derived using ratio change.

GSPREE and SPREE comparisons

We compared local authority-level ethnic group proportions produced using SPREE and GSPREE with the corresponding Census 2021 proportions and found GSPREE consistently performed better.

Across local authority and ethnic group comparisons, GSPREE was closer to the corresponding Census 2021 proportion more often than SPREE for all ethnic groups.

Table 2: GSPREE was closer to the corresponding Census 2021 proportion more often than SPREE for all ethnic groups, across local authorities

Number of local authorities where SPREE and GSPREE proportions were closer to Census 2021 proportions, by ethnic group, England and Wales, 2021

Ethnic Group	GSPREE	SPREE	Tie
Asian	246	83	1
Black	198	127	5
Mixed	161	158	11
Other	256	66	8
White	221	107	2
Total	1082	541	27

Source: Office for National Statistics

Notes

1. SPREE is Structure Preserving Estimator, GSPREE is Generalised Structure Preserving Estimator.
2. Data for Cornwall and the Isles of Scilly were combined for this analysis; therefore 330 local authorities are included.
3. Counts are based on comparisons of Census 2021, GSPREE and SPREE proportions rounded to two decimal places.

The largest differences were seen for the Asian and Other ethnic groups, where GSPREE was closer to Census 2021 proportions in substantially more local authorities than SPREE. The Mixed ethnic group showed the smallest difference between methods, with both models producing similar results across many local authorities. Overall, the findings provide evidence that combining administrative data, survey data and estimation methods can produce ethnicity estimates that are closer to Census 2021 than using administrative data alone.

Comparison of GSPREE estimates with Census 2021

As GSPREE showed closer agreement with Census 2021 than SPREE across local authorities, the following analysis focuses on GSPREE estimates. Differences between GSPREE and Census 2021 proportions were generally small, although patterns varied by ethnic group and local authority.

Figure 3: The GSPREE method produces estimates that are generally close to Census 2021 across local authorities in England and Wales

Number of local authorities where GSPREE ethnic group proportions were above, below or equal to Census 2021 proportions, by ethnic group, 2021

Notes:

1. GSPREE is Generalised Structure Preserving Estimator.

2. Local authority boundaries are as of 2021.

3. Data for Isles of Scilly have been combined with Cornwall.

Downloac

Figure 3 shows that differences between GSPREE and Census 2021 proportions were generally clustered around zero for all five ethnic groups across local authorities. A small number of local authorities showed larger differences, particularly for the Mixed, Other and White ethnic groups.

To provide further information on whether GSPREE estimates were higher or lower than the corresponding Census 2021 proportions, Table 3 shows the number of local authorities where GSPREE estimates were above, below or equal to Census 2021.

Table 3: GSPREE estimates were more often above Census 2021 proportions for Black, Mixed and White ethnic groups, and below for Asian and Other ethnic groups
Number of local authorities where GSPREE ethnic group proportions were above, below or equal to Census 2021 proportions, by ethnic group, 2021

Ethnic group	Over Census	Under Census	Equal
Asian	146	177	7
Black	173	146	11
Mixed	227	101	2
Other	49	276	5
White	204	122	4

Source: Office for National Statistics

Notes

1. GSPREE is Generalised Structure Preserving Estimator.

2. Data for Cornwall and the Isles of Scilly were combined for this analysis; therefore 330 local authorities are included.

3. Counts are based on comparisons of Census 2021, GSPREE and SPREE proportions rounded to two decimal places.

GSPREE estimates were more often above Census 2021 proportions for the Black, Mixed and White ethnic groups, and more often below for the Asian and Other ethnic groups. For the Other ethnic group, estimates were below Census 2021 proportions in 276 local authorities and above in 49.

Greater variation was seen for the Mixed and Other ethnic groups. For the Other ethnic group, estimates tended to be lower than the corresponding Census 2021 proportions, indicating underestimation across local authorities. One of the most notable examples was for Enfield, where the Census 2021 proportion for the Other ethnic group was 12.14% compared with a GSPREE estimate of 4.9%.

For the Mixed ethnic group, most local authorities remained relatively close to the Census 2021 proportions, but a small number of local authorities showed substantially larger differences. For example, in East Lindsey, the Census 2021 Mixed ethnic group proportion was 1.04% compared with a GSPREE estimate of 4.92%, a difference of 3.88 percentage points. Similar patterns were observed in Mid Suffolk (1.42% compared with 5.17%) and Ipswich (4.59% compared with 8.39%).

7 . Future developments

The findings from this research suggest that it is feasible to produce population estimates by ethnic group of reasonable quality using the ratio change and Generalised Structure Preserving Estimator (GSPREE) methods. These findings are based on a combination of the survey and administrative data currently available.

Differences between the estimates and Census 2021 are expected, as no alternative data source or estimation method can fully replicate the coverage and quality of a census. The main consideration is therefore the size of these differences, whether they can be explained, and whether they are acceptable for users' intended uses.

The results presented in this article provide a proof of concept. Before deciding whether to implement the methods in the production of official statistics on the population by ethnicity, further work is needed to engage with users and continue testing and development.

Investigate more detailed ethnicity and demographic breakdowns

We will explore the feasibility of producing estimates for more detailed ethnic group classifications, as well as additional breakdowns by characteristics such as age and sex.

Develop measures of uncertainty

We plan to investigate the feasibility of producing measures of uncertainty for estimates to help users better understand the reliability of the results and how to interpret differences over time or between areas or ethnic groups.

Explore variations in implementation of ratio change

We plan to investigate the potential of using individual administrative data sources, rather than the Admin-Based Ethnicity Dataset (ABED), to calculate the year-on-year changes to be used in the ratio change method. If it is feasible to use single administrative data sources for subgroups of the population, this would remove the introduction of quality issues arising from when sources are linked.

Explore variations in implementation of GSPREE

We plan to test whether we can improve the quality of local authority estimates through applying separate GSPREE models for each region. We will also compare regional estimates produced using ratio change with estimates produced by summing the local authority estimates produced using GSPREE.

Re-run estimates using updated data inputs

When available, we will re-run our testing with new versions of ABED and with Annual Population Survey (APS) data that have been reweighted to account for Census 2021.

Produce estimates for additional reference years

Future work will extend the analysis from 2021 to additional years, allowing us to assess how the methods perform over time. Producing estimates for multiple years will also support the development of an annual time series of population estimates by ethnic group for England and Wales.

User engagement

User engagement will continue to play an important role in the development of ethnicity statistics. This research builds on engagement undertaken through programmes including Census 2021 and the Future of Population and Migration Statistics (FPMS), which has helped shape our understanding of user needs and priorities.

We plan to undertake some further user engagement exercises over the coming months to gather feedback on our research presented and inform our next steps. We will also draw on feedback from the Government Statistical Service Harmonisation Consultation on user needs for additional response options in a future ethnicity standard.

8 . Provide feedback

We invite feedback on:

- the methods tested
- how ethnicity estimates would be used
- quality requirements for different uses
- priorities for breakdowns, including ethnic group, geographic, age and sex

Your feedback will help us continue to develop ethnicity statistics that meet user needs and support a wide range of uses. If you would like to share feedback or would be interested in taking part in future user engagement activities related to this work, please contact us at pop.info@ons.gov.uk.

9 . Definitions

Ethnic group

For this article, ethnicity is based on how individuals describe their own ethnic background and cultural heritage.

A harmonised standard is available for collecting data on ethnicity. The standard is currently under review to ensure that it continues to meet user needs and reflects the diversity of the population. The Government Statistical Service (GSS) will publish an update on 12 November 2026. Harmonisation means using comparable and consistent questions, definitions and data standards across different data collections so that statistics can be compared more easily and consistently.

Since ethnicity is a multifaceted and changing phenomenon, different approaches to measuring ethnic groups are available and have been used over time. These include where you were born, heritage, culture, national identity, religion, skin colour and language. What seems to be generally accepted, however, is that ethnicity includes all these aspects, and others, in combination.

When providing their ethnicity in administrative data, surveys or the census, respondents select from the response options available to them or choose not to answer the question. Some data collections, including the census, may also allow respondents to provide a more detailed response using a write-in option. The response options available vary across data collections. Some administrative data sources use more detailed response categories, which we have aggregated into the ethnic groups used in this article.

Statistical Population Dataset (SPD)

An SPD is tabular data that aim to capture the usually resident population on a reference date (30 June) for a given year within England and Wales. It is constructed from administrative sources that are linked together, and the combined evidence is used to derive who is usually resident and their characteristics of age, sex and geographic location.

10 . Data on developing population estimates by ethnicity

Developing population estimates by ethnicity for England and Wales
Dataset | Released 6 November 2026
Data for research into methods for producing annual population estimates by ethnic group and results from testing against Census 2021.

11 . Related links

[A summary of data and publications related to ethnicity]
Dataset | Released 15 January 2026
This document provides links to datasets and publications related to ethnicity statistics.

[Developing admin-based ethnicity statistics for England and Wales: 2020]
Article | Released 7 February 2023
Research update on producing population statistics by ethnic group for England and Wales from administrative data, with comparisons with Census 2021 estimates.

[Developing admin-based ethnicity statistics for England: 2016]
Article | Released 23 May 2022
Update on feasibility research on producing statistics on the population by ethnic group for England from administrative data, following the inclusion of additional data sources and method changes. These research outputs are not official statistics.

[Change over time in admin-based ethnicity statistics, England: 2016 to 2020]
Article| Released 23 May 2022
Feasibility research on producing statistics on the population by ethnic group for England for 2016 to 2020 from administrative data, exploring change over time. This builds on work to improve the 2016 admin-based ethnicity statistics. These research outputs are not official statistics.

12 . Cite this page

Office for National Statistics (ONS), released 6 November 2026, ONS website, article, Developing population estimates by ethnicity for England and Wales