

Article

# Quality of Census 2021 gender identity data

A final summary of our recent investigations into the quality of census data from the gender identity question and further planned research into this topic.

Contact:  
Census customer services  
census.customerservices@ons.  
gov.uk  
+44 1329 444972

Release date:  
8 November 2023

Next release:  
To be announced

## Notice

### 9 November 2023

We have made a minor text change to provide clarity around one of the figures in section 5.

# Table of contents

1. [Main points](#)
2. [Overview](#)
3. [Gender identity and the 2021 Census](#)
4. [Collecting and processing the data](#)
5. [Investigating respondent error](#)
6. [Consistency of census estimates with other sources](#)
7. [Engagement with users of the data](#)
8. [Release of additional supporting information](#)
9. [Further research into gender identity](#)
10. [Data](#)
11. [Glossary](#)
12. [Related links](#)
13. [Cite this article](#)

# 1 . Main points

- Census 2021 showed that 262,000 people (0.5% of the population aged 16 years and over) reported that their gender identity was different to their sex registered at birth.
- Census estimates on gender identity are broadly consistent with the best available comparator of the GP Patient Survey and international comparators.
- There are some patterns in the data that are consistent with, but do not conclusively demonstrate, some respondents not interpreting the question as intended; given other sources of uncertainty, not least the impact of question non-response, we cannot say with certainty whether the census estimates are more likely to be an overestimate or an underestimate of the total number of trans people aged over 16 years in England and Wales.
- As recognised in the Office for Statistics Regulation [interim report](#) on their review of these data, there are good reasons to expect higher levels of uncertainty with the estimates relating to the gender identity question than for other census topics.
- Users of the data have told us that they understand the relatively high levels of uncertainty in the estimates for this topic and that they would welcome more information to support their use of the data, therefore we have released additional tables alongside this report.
- We outline our plans for further work on gender identity statistics, recognising that while the census estimates are an important step forward in the evidence base for analysis on this topic, more work is needed to develop gender identity statistics for which the quality and extent of uncertainty is better understood.

## 2 . Overview

Census 2021 was the first census in England and Wales to ask about people's gender identity (see [Section 11: Glossary](#)). Users from both central and local government as well as charity and voluntary sectors, noted the importance of collecting high-quality data about gender identity to meet the requirements of the [Equality Act 2010](#) and the associated [Public sector equality duty](#).

The voluntary question went through a [rigorous development and testing process](#) which involved trans and non-trans people and, through the Census Rehearsal, people who did not have English as their main language. The final question was evaluated as meeting the requirements of public acceptability, being understood by respondents and providing the data needed by users.

The gender identity question appeared in the English (and Welsh) individual questionnaire as follows.

Is the gender you identify with the same as your sex registered at birth?

This question is voluntary:

- Yes
- No, enter gender identity (respondents were shown a write-in box)

As with all census topics, we carried out quality assurance on the data for gender identity before publishing the results. The published results were accompanied by [information on the quality of the data](#). Our analysis of the [capture and coding of the gender identity question](#) confirms that we have accurately reported what people told us in the census. Since that report, in response to feedback from some users, we have:

- extended our quality assurance of the gender identity data by carrying out additional research into how the data were collected and processed
- investigated whether there is any evidence of respondent error distorting the census estimates
- reviewed how consistent the census results are with other sources of information
- considered how well the information we have published meets the needs of main users of these data

The Office for Statistics Regulation published an interim review on [statistics on gender identity based on data collected as part of the 2021 England and Wales Census](#). One of its main recommendations was for the Office for National Statistics (ONS) to publish their plan for analysis of gender identity statistics from the census within four weeks of their report. We consider that it would be in the public interest to instead publish the results of our recent investigations into these statistics as soon as possible, so this report presents these results and outlines next steps for developing gender identity statistics as part of our [development of population statistics](#).

### 3 . Gender identity and the 2021 Census

We published the first census results for this topic in [Gender identity, England and Wales: Census 2021](#) with more detailed data being made available in January 2023 and April 2023. These releases were accompanied by [quality information](#) to help in understanding the data. Links to the releases are provided in [Section 12: Related links](#).

The census gender identity question was a voluntary question that was answered by 94% of people aged 16 years or over who responded to the census. The census reported that 262,000 people (0.5% of the adult population) in England and Wales had a gender identity different to their sex registered at birth (described in this report as being trans).

As recognised in the [Office for Statistics Regulation interim report](#) on their review of these data, there are good reasons to expect higher levels of [uncertainty](#) in the estimates relating to the trans population than for other census topics. As well as it being a voluntary question, gender identity is a sensitive topic and one which relies on self-identification of a concept which people may interpret differently. Additionally, those completing the census on behalf of others may not report, or even know, how a person identifies.

The trans group of 0.5%, or 1 in 200, is a small proportion of the population. Where numbers are based on a population of a small size, they have a higher risk of being impacted by errors, such as respondent error ([see Section 5: Investigating respondent error](#)). This also means that the final estimates could be substantially affected by bias in the census respondents who decided not to answer this voluntary question (more than 10 times the number of people who recorded that they were trans). Finally, there are limited sources collecting gender identity data to compare and check census results with.

More detailed results on gender identity indicated that the trans population is not spread equally across all groups of the population. For example, 1.00% of people aged 16 to 24 years recorded that they were trans compared with 0.54% across the population as a whole. While this difference by age group could be expected there were other patterns in the data that could be considered unexpected. For example:

- the trans population made up 0.3% in the White: English, Welsh, Scottish, Northern Irish or British ethnic group, compared with 1.6% of people in the Black, Black British, Black Welsh, Caribbean or African ethnic group
- 0.4% of the population who had English (English or Welsh in Wales) as their main language reported as trans compared with 1.6% of people who had a different main language
- among people whose main language was not English (English or Welsh in Wales), the trans population made up 1.0% of people who spoke English very well compared with 2.1% of people who did not speak English at all
- some areas with large ethnic minority populations had a higher rate of trans responses than might be expected, for example, Newham and Brent are local authorities with over 1% of respondents identifying as having a gender identity different from their sex registered at birth. Since ethnicity can be associated with English language proficiency, this suggests that some respondents whose first language is not English or Welsh did not understand the question as intended

While these patterns may be unexpected, this does not mean they are necessarily wrong. Some local authorities may attract trans people because of established communities, and unexpected patterns can result from correlations between variables.

For example, as the trans population is relatively concentrated in younger age groups, that might naturally result in a similar concentration in an ethnic group with a similarly young age distribution ([Section 5: Investigating respondent error](#) explores how different groups of people answered the question while accounting for this potential correlation between age and other variables). Furthermore, any expectations of how the transgender population varies over different population subgroups were built on limited evidence — precisely the evidence gap that inclusion of this question on the census was intended to fill.

## 4 . Collecting and processing the data

A detailed account of the work summarised in this section is provided in [Collecting and processing data on gender identity, England and Wales: Census 2021](#).

The question was voluntary and was answered by 94% of people aged 16 years and over responding to the census. As with the other voluntary questions on the census we did not impute "missing" answers to this question and people who decided not to answer the question are shown in a "not answered" category in published statistics.

Coding write-in responses (that is, assigning the text that a respondent wrote in as their gender identity to one of the categories used in published statistics) to this question was complex but uncertainties and small inconsistencies in coding do not have a practical effect on the use of the data. Write-in responses that did not refer to a gender identity were disregarded and treated as if the question had not been answered.

Statistical methods used to adjust the census data for undercoverage — that is, people not included on a census return — have been applied correctly. A method used to improve the accuracy of census estimates of students will have slightly lowered the estimate of the trans population, but this does not have a substantial impact on the usefulness of the data.

In summary, this strand of our research into the collection and processing of the census data for this topic provided no evidence that the design of the question or the statistical processing of the collected data had an adverse effect on the quality of the published statistics.

## 5 . Investigating respondent error

We conducted further investigations into whether there was any evidence of respondent error having an impact on the census estimates.

Table 1: Proficiency in English by gender identity  
England and Wales, all residents aged 16 years and over, 2021

| England and Wales                                               | Gender identity: the same as sex at birth | Gender identity: different from sex registered at birth, but no specific identity given | Trans woman | Trans man | Non-binary | All other gender identities | Not answered |
|-----------------------------------------------------------------|-------------------------------------------|-----------------------------------------------------------------------------------------|-------------|-----------|------------|-----------------------------|--------------|
| <b>Main language is English (English or Welsh in Wales)</b>     | 41,296,005                                | 69,165                                                                                  | 35,815      | 35,735    | 29,055     | 16,020                      | 2,428,405    |
| <b>Main language is not English (English or Welsh in Wales)</b> | 4,093,630                                 | 48,610                                                                                  | 11,760      | 12,700    | 1,205      | 2,055                       | 486,220      |
| <b>Of which, can speak English</b>                              |                                           |                                                                                         |             |           |            |                             |              |
| <b>Very well</b>                                                | 1,786,555                                 | 9,930                                                                                   | 3,625       | 4,115     | 1,075      | 850                         | 164,670      |
| <b>Well</b>                                                     | 1,480,955                                 | 23,295                                                                                  | 5,150       | 5,785     | 110        | 675                         | 202,060      |
| <b>Not well</b>                                                 | 703,400                                   | 13,135                                                                                  | 2,575       | 2,495     | 15         | 430                         | 101,230      |
| <b>Not at all</b>                                               | 122,725                                   | 2,245                                                                                   | 410         | 300       | -          | 95                          | 18,255       |

Source: Census 2021 from the Office for National Statistics

### Notes

1. For the purposes of this table counts lower than 10 have been suppressed and shown as '-' and counts have been rounded to the nearest 5. Figures may therefore be slightly different to estimates provided in standard outputs.
2. Totals have been rounded independently and will not necessarily be the sum of the shown components.

We investigated patterns of trans identification by ethnic group, country of birth and proficiency in English. There were clear patterns of trans identification being higher for people born outside the UK and people with lower proficiency in English (as shown in Table 1). Of those who indicated their gender identity was the same as their sex registered at birth, around 2% did not speak English well. This proportion was around 13% for those who indicated their gender identity was different to their sex registered at birth and did not write-in a specific gender identity. These patterns were still evident after adjusting for the different age profiles of the groups involved.

These patterns might be thought consistent with some respondents not interpreting the question as we had intended but could also be affected by other considerations such as cultural factors. For example, it is possible (but difficult to confirm) that trans migrants might have specifically chosen the UK because of its civil rights legislation and greater social acceptance than many other countries, impacting the trans proportion among that population group. The Foreign Office and National LGBT Survey (although self-selecting) published information supporting this, please see [Section 12: Related links](#).

Some further evidence of the possibility of respondent error was provided by examining how the recorded trans population answered the gender identity question. Respondents recorded as trans who were born outside the UK or had a lower proficiency in English were around twice as likely to tick the box and not write in a response compared with other trans respondents. This could be attributed to either the tick-box being answered inaccurately, or those with lower proficiency in English not being sure how to accurately describe their identity in English, or simply wanting to minimise the amount they needed to write in their response to this voluntary question.

A potentially helpful source of information about the quality of the census is the Census Quality Survey (CQS). This is a survey conducted shortly after the census to assess the potential level of respondent error for each census question. The results of the survey were published in our March 2023 release and showed high overall agreement between how people answered the gender identity question on the census and in the survey. More detailed analysis for subgroups of the sample suggest that agreement rates were much lower for people who recorded they were trans on the census. However, this analysis comes with several important caveats:

- it is based on a small subset of the CQS sample – fewer than 300 people who recorded that they were trans on the census
- the CQS collected information via telephone, so there is a risk of others in a household overhearing a person's responses, or a respondent may feel uncomfortable sharing this information over the telephone; therefore, they may be less likely to respond to sensitive questions such as gender identity, resulting in potential non-response bias to the CQS
- analysis also shows limited evidence (based on very small sample sizes) that the CQS may not have worked well in correctly capturing the answer to this sensitive question from trans people who did respond to the CQS

While we cannot draw any confident conclusions from this analysis, we will take forward further analysis of the CQS data as part of the development of the harmonised question on gender identity.

To summarise our investigation into potential respondent error: there are patterns in the data that are consistent with some respondents not interpreting the question as we had intended. These patterns reflect those respondents with lower proficiency in English and characteristics likely associated with English language proficiency such as ethnicity and country of birth. However, for the reasons outlined in this report, we cannot say whether the census estimates are likely to be an overestimate or an underestimate of the true value, given other sources of uncertainty, not least the potential impact of question non-response. Therefore, the overall impact on the data of any misinterpretation of the question cannot be determined. To assist users in interpreting the detailed breakdowns of the data, we have produced additional supporting tables (see [Section 8: Release of additional supporting information](#)).

## 6 . Consistency of census estimates with other sources

Comparing census estimates with other sources of data is a standard part of our quality assurance for every census topic. This work was limited for the gender identity topic as there were no previous census data and few other sources of statistical information.

We compared the census estimates for England with the results of the [2021 and 2022 GP Patient Survey](#) (GPPS) conducted for NHS England. This showed the results were broadly consistent between the two sources.

Table 2: Trans population as percentage of all people aged 16 years and above

| <b>Census 2021<br/>(England and Wales)</b> | <b>Census 2021<br/>(England only)</b> | <b>2021 GP Patient<br/>Survey (England only)</b> | <b>2022 GP Patient<br/>Survey (England only)</b> |
|--------------------------------------------|---------------------------------------|--------------------------------------------------|--------------------------------------------------|
| 0.54%                                      | 0.55%                                 | 0.65%                                            | 0.76%                                            |

Source: Census 2021 from the Office for National Statistics; GP Patient Survey 2021 and 2022 from NHS England.

Note: Analysis of GPPS data in this article has been conducted by the ONS. We are engaging with NHS England to help us understand patterns in both data sources.

The census estimate for the percentage of the English population who are trans (0.55%) is slightly lower than the weighted 2021 GPPS estimate (0.65%). The patterns by age, ethnic group and religion identified in the census data also appear in the GPPS estimates.

The GPPS is a good comparator for the census data in that it has a very good coverage of the population and collected data for a similar point in time. It has a drawback as a comparator in that the wording of the question is very similar to that on the census and there could be a risk that any respondent errors are correlated across the two sources.

Please note that although the questions are similar, they are worded slightly differently with different response options, and are positioned differently in the two questionnaires, which may impact on how people respond to them. In addition, although the GPPS has a good coverage of the population, the census and GPPS draw from different sample frames.

We also compared the census estimates of the trans population with the 2021 Canadian Census and surveys in the United States. The [Canadian 2021 Census results](#) show that 0.33% of Canada's population provided an answer to the gender question that was different to their answer to the sex question. In comparison, the England and Wales census revealed that 0.54% of the population identify as a gender that is different to their sex registered at birth.

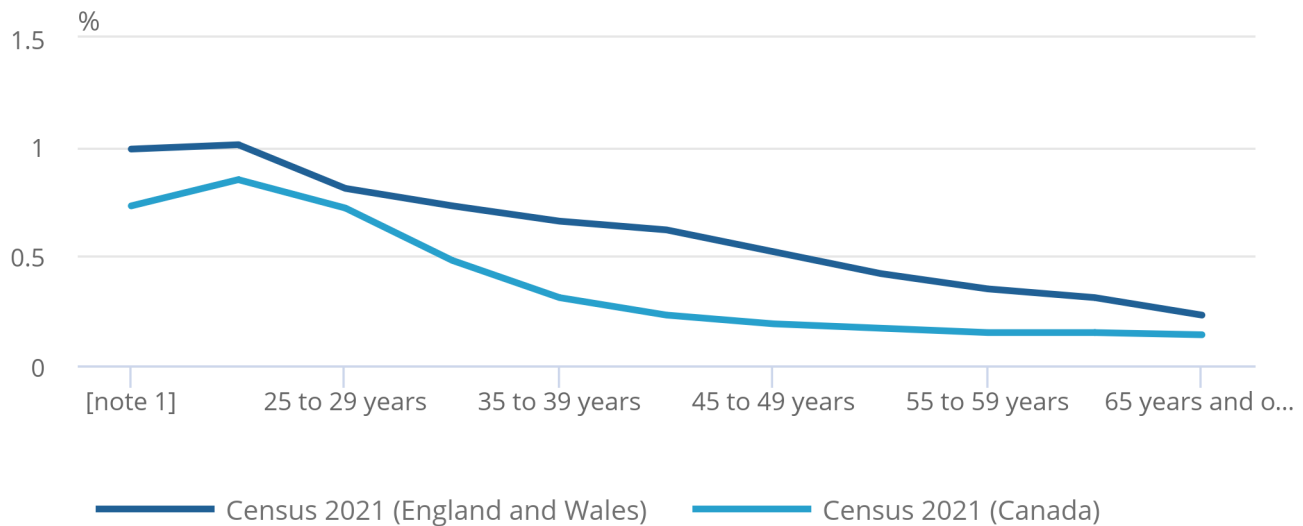


## Figure 1: Trans population as a percentage of each age group

England and Wales, and Canada, 2021

### Figure 1: Trans population as a percentage of each age group

England and Wales, and Canada, 2021



Source: Census 2021 from the Office for National Statistics; 2021 Census of Population from Statistics Canada

#### Notes:

1. The “trans” figure for Canada relates to those who provided a response to the gender question that was different to their response to the sex question.
2. The first age group [note 1] relates to anyone aged 16 to 19 years in England and Wales Census 2021 data and anyone aged 15 to 19 years in Canada’s Census 2021 data.

Differences between the results on this topic from the two censuses might reflect genuine differences between the populations of Canada and of England and Wales, or could be because of differences in the question design and data collection. While these comparisons are difficult to interpret, they provided no evidence to suggest that the England and Wales census estimates were incorrect.

Studies in the United States suggest that between 0.3% and 0.6% of the country's population identify as a gender that is different to their sex registered at birth. This aligns with the findings from the England and Wales census, which estimated that 0.54% of the population identify as a gender that is different to their sex registered at birth. More information on the United States sources can be found in [Section 12: Related links](#).

Additionally, some users suggested examination of the census results with the decennial [National Survey of Sexual Attitudes and Lifestyles](#) (NATSAL-4) that was scheduled to include a gender identity question but because of the coronavirus (COVID-19) pandemic, this survey was delayed. While a question on gender identity was included in [their COVID study in 2020 and 2021](#), the question phrasing differed to Census 2021 and asked respondents to describe how they think of themselves. We considered that the sample size (6,654) and return rate (4%) was too small to draw meaningful comparisons.

In summary, this strand of our research comparing the census estimates with other available sources found no reason to conclude that the census statistics on gender identity were implausible.

## 7 . Engagement with users of the data

An important aspect of the quality of these data is whether the statistics meet user needs, and how well the information we have published to support the statistics assists users in meeting those needs.

To understand this better, we have sought the views of over 20 organisations, from government departments, local authorities, academia, and the non-profit and commercial sector. They have told us:

- they expect higher uncertainty around census estimates for this topic, particularly as it is a voluntary question
- they are not solely relying on the census estimates in isolation for their purposes but view the census results as an important step in improving the availability and understanding of data on the gender identity topic
- they would welcome additional information to help them understand the census data for this topic and its associated uncertainties, and thus how to use the data appropriately
- information on write-in responses would be useful to understand the language used by those who identify as a gender identity different to their sex registered at birth.

## 8 . Release of additional supporting information

As with every census topic, we have published a [quality information page for gender identity and sexual orientation](#). This page notes the higher level of uncertainty around census results on gender identity and some particular quality-related aspects of the data, which should be considered when using the data.

We are also publishing additional information alongside this report to help people to have a greater understanding of the data. This additional information has four parts.

### Proportion of write-in responses for local authorities

While we cannot quantify potential respondent error in the estimates of the trans population, it is reasonable to suggest that someone is less likely to have mis-reported themselves as having a gender identity different to their sex at birth if they ticked the relevant box and also wrote in their gender identity. This file contains the proportions of the estimated trans population within each local authority who provided a write-in response.

These do not provide alternative estimates of the trans population, since many trans people will have accurately ticked the box and not provided a write-in response. However, they can be used as broad indicators of which local authorities are likely to have higher certainty in their estimates of the trans population. If gender identity estimates for a particular local authority are based on a low proportion of write-in responses (therefore a higher proportion of tick-box only responses), then users may wish to exercise greater caution in using these statistics and consider them alongside other data sources and local knowledge for practical purposes.

### Topic distributions by gender identity

These tables are an alternative presentation of the data, at an England and Wales level, that have already been published. For each topic, they show the distribution across categories for the estimated trans population and additionally the distribution based only on those who provided a write-in response to the question. Topics where both distributions are similar are likely to be relatively unaffected by any respondent error resulting from simply ticking the wrong box. A greater degree of care would be advisable where the distributions are markedly different.

### Write-in responses to the gender identity question

This file contains the counts and associated coded category for every interpretable write-in response (with at least 10 occurrences). This information allows users to form their own assessment of the coding and to understand the contribution made to the final estimates by particular write-in responses.

## Additional tables on gender identity

These tables go beyond our previous data releases and provide additional analyses identified during our discussions with users as being valuable in understanding the data.

## 9 . Further research into gender identity

We are continuing to publish analysis and data from Census 2021 England and Wales. Information on our release plans is provided through our [Census page](#) and the [Release calendar](#).

The publication of this report and supporting information concludes our planned research directly on the quality of the census data on gender identity.

We will consider the Office for Statistics Regulation's final report from their review into these statistics with a view to publishing any further information recommended in that report.

We will continue our discussions with users when reviewing the gender identity questions, related questions on sex, and in our future transformation work. If users would like to request additional census data for their own investigations, they can request this via our [commissioned table service](#) subject to statistical disclosure considerations.

Further research on collection of data on gender identity in both surveys and administrative data sources will be part of the Office for National Statistics's (ONS's) ambitious and radical programme to develop population and migration statistics. The ONS's recent public consultation on the future of these statistics has received over 700 responses representing a wide range of views from a variety of organisations and individuals, which will inform recommendations by the National Statistician.

The consultation will inform our work on developing a harmonised standard for collecting data on gender identity. This work will also draw on further analysis of the results of the Census Quality Survey, discussed in [Section 5: Investigating respondent error](#).

We welcome any comments or suggestions on the investigations presented here, or on additional information that would help in using the estimates. Please email [census.customerservices@ons.gov.uk](mailto:census.customerservices@ons.gov.uk).

## 10 . Data

[Census 2021: Quality of Census 2021 gender identity data](#)

Dataset | Released 8 November 2023

This dataset provides supporting information for the Quality of Census 2021 gender identity data release.

## 11 . Glossary

### Gender identity

Gender identity refers to a person's sense of their own gender, whether male, female, or another category such as non-binary. This may or may not be the same as their sex registered at birth.

### Trans

The term "trans" is used in this report to describe anyone who stated in the census that their gender identity was different to their sex registered at birth. This includes people who identify as a trans man, trans woman, non-binary or with another minority gender identity.

## **Non-binary**

Someone who is non-binary does not identify with the binary categories of man and woman. In these results the category includes people who identified with the specific term "non-binary" or variants thereon. However, those who used other terms to describe an identity that was neither specifically man nor woman have been classed in "All other gender identities".

## **Trans man**

A trans man is someone who was registered female at birth, but now identifies as a man.

## **Trans woman**

A trans woman is someone who was registered male at birth, but now identifies as a woman.

## **Sex**

This is the sex recorded by the person completing the census. The options were "Female" and "Male".

## 12 . Related links

### [Collecting and processing data on gender identity, England and Wales: Census 2021](#)

Methodology | Published 19 June 2023

Methodology for collecting and processing data on gender identity in Census 2021.

### [Gender identity by ethnic group](#)

Dataset | Published 6 April 2023

Census 2021 estimates that classify usual residents aged 16 years and over in England and Wales by gender identity and ethnic group.

### [Gender identity by age and sex, England and Wales: Census 2021](#)

Article | Published 25 January 2023

Gender identity by age and sex of usual residents aged 16 years and over in England and Wales, Census 2021 data.

### [Sex and gender identity question development for Census 2021](#)

Methodology

How we researched, developed and tested the Census 2021 questions on sex and gender identity.

### [Census 2021: Final guidance for the question "What is your sex?"](#)

Article

How we evaluated the results of our research and testing to finalise the Census 2021 online help guidance for the question "What is your sex?".

### [Lesbian, gay, bisexual and transgender: foreign travel advice](#)

Foreign and Commonwealth Office | Last updated 19 June 2023

Foreign, Commonwealth and Development Office (FCDO) advice for lesbian, gay, bisexual and transgender (LGBT+) people travelling abroad.

### [National LGBT Survey Summary Report](#) (PDF, 2.55MB)

Government Equalities Office | July 2018

Main findings from the 2017 nationwide LGBT survey.

### [Canada is the first country to provide census data on transgender and non-binary people](#) (PDF, 588KB)

Report | Published April 2022

Summary of the gender results from Census 2021 in Canada.

### [How Many Adults and Youth Identify As Transgender in The United States](#) (PDF, 3.75 MB)

Report | Published June 2022

Results from analysis of the Behavioral Risk Factor Surveillance System indicating how large the potential transgender population is in the United States.

### [Transgender Demographics: A Household Probability Sample of US Adults, 2014](#) (PDF, 584KB)

Report | Published February 2017

Results from analysis of the Behavioral Risk Factor Surveillance System indicating how large the potential transgender population is in the United States.

### [Transgender Population Size in the United States: a Meta-Regression of Population-Based Probability Samples](#) (PDF, 988KB)

Report | Published February 2017

Results from analysis of multiple different population surveys indicating how large the potential transgender population is in the United States.

## 13 . Cite this article

Office for National Statistics (ONS), released 8 November 2023, ONS website, article, [Quality of Census 2021 gender identity data](#)