



Office for  
Statistics Regulation



UK Statistics  
Authority

# Collecting and reporting data about sex and gender identity in official statistics

## A guide for official statistics producers

Updated 17<sup>th</sup> December 2024

# Contents

Key Messages	3
Important Information	4
About this guidance	4
How this guidance should be used	4
Terminology used in this guidance	5
OSR programme of work on sex and gender identity	6
Questions used to collect data about sex and gender identity	7
Proportionality	8
Legal requirements	9
Additional information	9
Deciding whether to collect	10
OSR's expectations	10
Questions for statistics producers to consider	11
Deciding how to collect	13
OSR's expectations	13
Questions for statistics producers to consider	14
How to report and considerations for onward use of the data	16
OSR's expectations	16
Onward use	17
Questions for statistics producers to consider	17
Understanding user needs	18
OSR's expectations	18
Questions for statistics producers to consider	18
Case studies	20
What's changed in OSR's guidance?	21

# Key Messages

This guidance document seeks to offer producers of official statistics a set of tools, based on the standards of the [Code of Practice for Statistics](#), to help them think through how they collect and report data about sex and gender identity.

**The key message of this guidance is that official statistics producers should have a strong understanding of their own data collections and how these meet the needs of their users, that producers should be clear about what information they are collecting and why, and explain this information as straightforwardly as possible both to respondents and data users.**

- **It is the responsibility of official statistics producers to determine whether there is a legitimate need to collect data about sex, gender identity, both or neither, based on identified user and respondent needs. Producers should be clear about what they are collecting and why.**
- **Data about sex and gender identity should be explained and defined for the purpose of a particular set of statistics, and terms, including gender, should not be used interchangeably or as a substitute for each other.**
- **Producers must have a strong understanding of their own data collections and explain the strengths and limitations of their data and statistics.**
- **Statistics should be consistent and coherent with related statistics and data where possible.**
- **Producers should be proportionate when applying the principles of this guidance to their data collections based on the nature of their statistics and the needs of their users.**

The content of this guidance is determined by our remit as the Office for Statistics Regulation; therefore, there are some areas that it does not cover:

- This guidance does not set out harmonised standards for sex or gender identity definitions.
- This guidance does not assume a particular question set is used to collect data about an individual's sex and gender identity. This guidance remains relevant in its application of the principles of the Code regardless of the questions used to collect these data.
- This guidance does not determine the requirements of relevant legislation or whether they are being adhered to.

This guidance updates and supersedes OSR's previous draft guidance on collecting and reporting data about sex in official statistics, published in 2021.

We recognise that this is an evolving area, and we will keep this guidance under review and update it when appropriate. We welcome feedback on this guidance from statistics producers and others at any time – please email [regulation@statistics.gov.uk](mailto:regulation@statistics.gov.uk).

# Important Information

## About this guidance

The Office for Statistics Regulation (OSR) is the independent regulator of official statistics produced in the UK. We work to ensure that data and statistics serve the public good and meet society's needs for information. Through the standards of Trustworthiness, Quality and Value set out in the [Code of Practice for Statistics](#) (the Code), we support producers to develop and deliver statistics that meet the needs of users and are accompanied by supporting information to aid the appropriate interpretation of the statistics.

This guidance is intended to support official statistics producers in applying the principles of the Code when making decisions about the collection and reporting of statistics and data about sex and gender identity. While the guidance is aimed at official statistics producers, others working with data and statistics, both inside and outside government, may find it helpful.

Official statistics producers regularly collect and publish information on sex. In recent years, some statistics producers have started, or are considering, collecting information on gender identity to address previous data gaps in the statistics or to allow data to be collected which recognise how a person identifies.

This is a complex area that is evolving in our society, and where public understanding may vary. In this context, statistics producers face challenges in their engagement and in making decisions about data collection and reporting. This is hindered by the absence of harmonised definitions of sex or gender identity for official statistics production in the UK, and the fact that survey respondents and data users may have their own interpretation of the terms being used.

All official statistics producers are responsible for ensuring that their statistics continue to be useful, easy to access, remain relevant and support understanding of important issues. Applying this guidance should not be seen as a one-off activity; rather, it should be continuously applied to support ongoing development and improvements.

We consider this guidance to be relevant regardless of whether the data are collected through surveys or administrative data systems. Both these collection types are considered within this guidance, and each has different considerations and complexities for statistics producers to take account of.

For further advice or additional support on the production of their statistics about sex and gender identity, producers should consult their Head of Profession for Statistics, Chief Statistician or Lead Official for Statistics.

## How this guidance should be used

Each section of this guidance addresses a typical stage in the production of official statistics. Within each section, we have set out our expectations for statistics producers and key questions that they should consider. While understanding the needs of users of statistics is presented last within this guidance, we consider this to be a central factor in each of the preceding stages.

This guidance is for individuals and teams who are directly involved in statistical production. Our expectation is that this guidance will be used to ensure that statistics producers have a strong understanding of what data they are collecting and how their data collections meet the needs of their users and respondents, and that statistical information is communicated clearly and transparently to statistics users.

Heads of Profession for Statistics, Chief Statisticians and Lead Officials play a key role in leading work in this area. These officials are responsible for advising statistical and analytical teams on statistical practices, making decisions on approaches and ensuring coherence across outputs whenever possible. We expect them to use and promote this guidance within their departments.

We encourage all statistics producers to engage with existing professional networks and to consider developing new networks in order to share best practice and learn from others.

## Terminology used in this guidance

The Code states that data sources should be based on definitions and concepts that are suitable approximations of what the statistics aim to measure, or that can be processed to become suitable for producing the statistics.

Currently there is no single definition of sex or gender identity within the official statistics community. We also recognise that respondents and users may have differing interpretations of these terms. Given this, it is vital that producers clearly define and explain any terminology that they are using. It is the responsibility of statistics producers, rather than OSR, to decide on and explain the definitions used within any data collections and statistical outputs.

This guidance does not set out harmonised standards for sex or gender identity definitions, as this is not within our remit as a regulator. We have, however, looked at how sex and gender identity have been defined within guidance published by other organisations. To provide a framework for using this guidance, we have set out our interpretation of these terms. We recognise that there will be differences in how familiar people are with the terms, that some people may prefer to use alternative terms than those set out below and that some people may feel that certain terms do not apply to them.

When we use the term 'sex' in this guidance, we are referring to a binary variable categorised as female or male. In the UK, an individual's legal sex is recorded at birth based on their biological characteristics. Individuals over the age of 18 can apply for a [Gender Recognition Certificate \(GRC\)](#), which enables them to change their legally recognised sex, including on their birth certificate. Where an individual has a GRC, their legally recognised sex may differ from their sex recorded at birth.

When we use the term 'gender' in this guidance, we are referring to the socially constructed characteristics and behaviours of men, women, boys and girls. The concept of gender is not fixed and can vary from society to society and change over time. We acknowledge that there are different interpretations of 'gender' and that the history of this term is complex. In addition, within official statistics data collections, the term 'gender' has often been used interchangeably with 'sex'.

When we use the term 'gender identity' in this guidance, we are referring to an individual's innate sense of their gender. Gender identity is not constrained to the binary male/man

and female/woman categories and includes others such as non-binary and gender fluid. For some individuals, their gender identity may be different from their sex recorded at birth, and/or from their legally recognised sex.

When we use the term ‘transgender’ in this guidance, we are referring to individuals whose gender identity is different from their sex recorded at birth. Some statistics producers are currently collecting, or may be considering collecting, information on an individual’s trans status or trans history. This guidance does not explicitly cover this approach to data collection. However, we consider that many of the principles of this guidance are still useful and applicable.

Producers should be aware that an individual can change their sex on some documents to reflect their gender identity without changing their legally recognised sex with a GRC. These documents include a driving licence, a passport, medical records, employment records and a bank account. Statistics producers should account for this when they are producing definitions and providing guidance to respondents on how they should answer questions on sex or gender identity.

## Using terminology interchangeably

We recognise that the terminology in this topic area is evolving, which can be a challenge for statistics producers. As a result, some producers may currently be using terminology that does not accurately reflect what they are trying to collect and report data on. This underlines the importance of clearly defining and explaining the terms being used, as well as not using terms interchangeably or as substitutions for each other.

Through our work, we have seen instances where there is a lack of consistency and clarity around the term ‘gender’, both in data collection and in statistical reporting. In some cases, it is not clear whether producers are using the term gender as a substitution for sex or gender identity. For example, producers may collect data about sex, yet report statistics on the basis of gender. Alternatively, producers may ask a question on gender with the intention of collecting data about gender identity, but only provide binary response options of male/man and female/woman. This conflation of terms lacks clarity for respondents and users regarding exactly what data are collected and may affect the quality of the data reported.

We recommend that statistics producers who collect and report ‘gender’ should consider what they mean by this term and whether it accurately reflects what they are trying to measure. This should be clearly communicated to respondents and users of the statistics.

## OSR programme of work on sex and gender identity

This work forms part of a broader OSR programme on data about sex and gender identity. As part of this programme, we undertook a review of gender identity in the 2021 England and Wales Census. We published an [interim report](#) on 9 October 2023 and [a final report](#) on 12 September 2024. The publication of this final report concluded our review.

Learning from new evidence in Scotland’s Census, the Office for National Statistics (ONS) [wrote to us](#) on 5 September 2024 to request that the gender identity estimates from Census 2021 should no longer be accredited official statistics and should instead be classified as [official statistics in development](#). We wrote to ONS to support its proposal.

Our [final report](#) sets out the details of our regulatory decision and shares our recommendations to ONS on the steps it must take to help users of the census gender identity statistics understand their strengths and limitations. In addition, we set out the development work we consider is required on the Government Statistical Service (GSS) gender identity harmonised standard. Our recommendations include that, as part of the ongoing development of the GSS harmonised standard for gender identity, ONS should carry out further testing of the question and that the GSS publish advice for statistics producers who are currently using the gender identity question and for those who are thinking about adopting the standard.

We found that the issues highlighted in our final report were unique to the statistics on gender identity and therefore that all other outputs from the Census 2021 in England and Wales are unaffected by the decision and remain designated as accredited official statistics.

## Questions used to collect data about sex and gender identity

It is important that the collection and reporting of data about sex and gender identity is based on users' needs for data and statistics. How those data are captured will depend on various factors, including the questions asked to survey respondents and the related guidance, or the data required for administrative systems. Different approaches can be taken in designing questions to collect these data, and – while we support producers taking account of comparability, data quality and respondent acceptability – ultimately it is the decision of the statistics producer or producer organisation as to exactly which questions and what guidance are used.

The Government Statistical Service (GSS) is responsible for providing information to producers on harmonised standards for data collection. These include definitions, survey questions and suggested presentations for users.

Acting on the recommendations set out in [the final report](#) of OSR's review of gender identity in the 2021 England and Wales Census, the GSS Harmonisation Team, based in the Office for National Statistics (ONS), has [commenced a new phase of work](#) towards agreed harmonised standards for sex and gender identity data collection. [Sex harmonisation guidance](#) and new [Gender Identity Data Harmonisation interim guidance](#) are available for producers of statistics who are currently using, or thinking about including, a gender identity question on their surveys or in their administrative data collection. OSR will continue to engage with the GSS Harmonisation Team and with statistics producers regarding this work.

The Scottish Government has published guidance for public bodies on the collection of data on sex and gender, [Sex, gender identity, trans status – data collection and publication: guidance](#), which includes a section on questions.

Our guidance does not assume a particular question set is used to collect data about an individual's sex and gender identity. This guidance remains relevant in its application of the principles of the Code regardless of the questions used to collect these data.

## Proportionality

### Proportionality in data collection

We recognise that the statistical landscape is complex, and that producers will be balancing competing priorities and available resources. We consider that the level of work undertaken to decide whether to make a change to the existing collection of data about sex, or to introduce a collection of gender identity data, should be proportionate to the importance of this information to the users of those statistics. This same principle applies to the collection of this information, where the burden imposed on those providing their data should be proportionate to the benefits.

For some statistical outputs, it will be important to capture information that would be unique to either sex or gender identity, for example, when statistics aim to understand the service provision or experiences of a particular group of individuals. In these instances, it is vital that producers make informed decisions about what to collect. For other statistical outputs, there may not be a clear need to collect a specific type of information and the decision may not affect the usability of the data. In either scenario, producers should ensure that they are using clear and considered definitions and that these are explained to respondents and users of the statistics wherever possible.

We acknowledge that it may not be possible for producers to apply all the principles of this guidance to their data collection. Producers should aim to balance considerations of best practice; data quality; user and respondent needs; the extent of possible system changes; and the required resource. Regardless of the extent to which producers are able to apply this guidance, producers should always be open and transparent about their practices, decision making and the level of understanding of what data are being captured and how. This transparency will ensure that users are supported in the interpretation and use of the resultant statistics.

Producers should be proportionate when applying the principles of this guidance to their data collections based on the nature of their statistics and the needs of their users.

### Proportionality in data use

For a large proportion of the population, an individual's sex and gender identity will be aligned. For a considerably smaller proportion of the population, an individual's gender identity is different from their sex as recorded at birth or their legally recognised sex.

Given the current varying approaches to collecting data about sex and gender identity across the statistical system, there will be times where producers or users are analysing data based on different approaches or on collection methods that did not have clearly defined terms or guidance for respondents. Users should also be aware that even if definitions and guidance have been provided to respondents, it is possible that respondents will have their own interpretations and may have answered questions differently. Datasets may therefore contain a combination of data about sex or gender identity.

In these circumstances, producers should seek to understand what data are actually being captured and evaluate the impact of the approach on the interpretation of the statistics. This information should be clearly communicated to users as well as any associated limitations.



Clear information on what has been collected is important where users are looking to conduct specific analysis. These users should be made aware of any limitations that may affect their analysis. Where key variables of interest are unlikely to vary significantly by sex or gender identity, producers and users of the data should have confidence in conducting high-level analysis and drawing conclusions.

## Legal requirements

The Code clearly sets out that all statutory obligations governing the collection of data, confidentiality, data sharing, data linking and release should be followed. Producers should familiarise themselves with any legal requirements and guidance that apply to the data and statistics they are collecting and publishing.

Where relevant, producers should ensure that they are adhering to the [Equality Act 2010](#), including the [Public Sector Equality Duty](#) for England, Scotland and Wales, and the [Section 75 duties](#) for Northern Ireland. Producers should be aware that sex and gender reassignment are [protected characteristics](#) as set out in the Equality Act 2010.

Producers should also ensure that they are adhering to the principles set out in the General Data Protection Regulation (GDPR), including [data minimisation](#). Producers should be aware that information about an individual's transgender status may be considered as [special category data](#) in some circumstances. Special category data must be treated with greater care than other types of personal data and can only be processed if one of the specific conditions in Article 9 of the UK GDPR is met (see the Information Commissioner's Office (ICO) guidance on [lawful basis on processing special category data](#)).

It is not within OSR's remit to determine the requirements of relevant legislation or whether they are being adhered to. Producers should consult their legal advisors and, where necessary, liaise directly with relevant bodies, such as the [Equality and Human Rights Commission \(EHRC\)](#), [Equality Commission for Northern Ireland](#) or the [ICO](#), in order to ensure that data collections are fully compliant with legislation.

## Additional information

In 2021, the [Inclusive Data Taskforce](#) published a [report](#) on how the government can be more inclusive in its data. This report includes recommendations on how best to make a step-change in the inclusivity of UK data and evidence. Many of these recommendations relate to the collection of information on sex and gender identity, either directly or indirectly. We encourage producers to familiarise themselves with the report and consider how they can contribute to the aims set out.

For support on any ethical considerations related to the use of data, we recommend producers visit the UK Statistics Authority's [Centre for Applied Data Ethics](#) website for further information.

# Deciding whether to collect

## OSR's expectations

The Code states that data sources should be based on definitions and concepts that are suitable approximations of what the statistics aim to measure. In addition, the collection, access, use and sharing of statistics and data should be ethical and for the public good.

It is the responsibility of official statistics producers to determine whether there is a legitimate need to collect data about sex, gender identity, both or neither. Producers should also familiarise themselves with any legal obligations relating to their data collections, such as the [Public Sector Equality Duty](#) for England, Scotland and Wales, and the [Section 75 duties](#) for Northern Ireland. If producers are unsure what information they should collect, they should seek advice from their Head of Profession for Statistics, Chief Statistician or Lead Official in the first instance.

Producers should consider the principles of data minimisation as set out in the General Data Protection Regulation (GDPR), namely that “Personal data shall be adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed”. This means that producers should collect information on an individual's sex or gender identity only if there is a need to. In addition, producers should consider collecting information on both sex and gender identity data only if there is a clear user or legal need that cannot be met by collecting either data item individually. This is especially relevant where collecting a combination of sex and gender identity data means that producers are directly or indirectly collecting information on an individual's transgender status, which may be considered special category data in some circumstances. See ‘Legal requirements’ for more information.

Statistics should be consistent and coherent with related statistics and data where possible. Producers should ensure that they are familiar with how producers of related statistics are collecting data about sex or gender identity. Where it is not possible to be coherent with related statistics and data, producers should give reasons for the deviation and explain any implications for use.

Coherence is especially important to consider for datasets that are, or could be, linked together. The use of inconsistent terms or data across such datasets has the potential to cause problems when linking them together. It is key that producers are clear on the terminology and definitions being used in their data collections to support accurate data linkage and subsequent analysis of the data.

When deciding whether to collect data about sex or gender identity, producers should consider the likely quality of the data that they will receive. For surveys, there are many factors that could impact the quality of the data collected, for example, question acceptability, low response rates, response biases, mode effects and small sample sizes. For administrative data, the quality of the data can be affected by factors such as whether the information is provided by the individual or answered on their behalf.

Producers should be aware that when trying to capture information on small sub-groups of the population false positives will have a greater impact. By false positive, we are referring to when someone reports that they are part of the target group when they are not. The

effect of these incorrect responses will be amplified in small groups, which may then affect the quality of the resulting statistics. If the data are unlikely to be of sufficient quality to meet those needs, producers should consider whether it is appropriate to collect these data.

There will be times when producers are not able to meet the needs of everyone who has an interest in the collection of data about sex or gender identity. One example is when the needs of the data users differ from the needs of the individuals providing their data, such as survey respondents. For example, a survey may be collecting data about sex, but some respondents may report that they would prefer to provide their gender identity. We acknowledge that these situations can be difficult for producers to navigate and that there may not always be a clear answer that will meet the needs of everyone involved. This may be especially true for statutory data collections where there is little ability to change the collection. In these instances, producers should clearly explain why certain information is or is not being collected and be transparent and open about their decision-making processes and the evidence and priorities used to inform their choices.

## Questions for statistics producers to consider

### Do you need to or want to report on sex?

While it is likely these data will already be collected, if you are looking to report on sex you should review how data are captured to meet your reporting requirements.

Ask yourself:

- Do you have a clear reason to collect information on sex?
- Are you clear on how you are defining sex?
- Does your definition of sex meet the needs of the different users of your statistics?
- Do you understand what data you can, or should, collect about an individual's sex to comply with relevant legislation?

### Do you need to or want to report on gender identity?

If you are looking to report on gender identity, it is likely that you will need to introduce a new collection or source these data. This may require amending your data collection or asking a new question. However, you should also consider if the information is already captured through other sources that you can use, for example, through data sharing agreements where appropriate.

Ask yourself:

- Do you have a clear reason to collect information about gender identity?
- Are you clear on what you mean by gender identity?
- Does your definition of gender identity meet the needs of the different users of your statistics?
- Have you considered relevant guidance and legislation relating to the collection of gender identity?
- Have you considered the likely quality of the data and determined that it will be sufficient to meet the needs of your users?

If you are considering collecting data about both sex and gender identity, you should ask yourself all of the questions set out above. In addition, you should consider whether you will be collecting special category data, and if so, you should evaluate what additional arrangements you will need to meet legislation requirements.

# Deciding how to collect

## OSR's expectations

The collection and reporting of statistics about sex and gender identity should support a legitimate public interest and be done in the least intrusive way. Producers should ensure that data are collected in a respectful way and should understand any risks to data quality or survey response when asking for sensitive information from a person.

Individuals providing their personal data may have their own understanding of the terms sex and gender identity. For this reason, it is important that, where possible, producers provide guidance to respondents and interviewers on how these questions should be completed as well as explanations for why certain information is required. This guidance will help ensure accuracy and coherence across the data collection. In particular, producers should avoid collecting data or reporting on gender without clear guidance and explanation of what is meant by this term, as it is often unclear whether the aim is to collect information on an individual's sex or their gender identity.

The privacy and identity of individuals must be protected at all times during data collection, storage, analysis and reporting. This includes being clear and open with individuals who are providing information about how their data will be used and how they will be protected. Producers must apply relevant security standards to keep data secure.

When designing data collection questions, there are many factors that producers should consider. Producers should avoid using response options that might cause confusion by ensuring that they undertake appropriate question testing of respondent understanding and acceptability. Producers should ensure that the response options align with the concept that they are aiming to collect data on. For example, we would expect a gender identity question to provide response options that reflect that gender identity is not a binary concept.

Producers should consider whether they should provide additional response options which allow respondents to not provide a response, for example, where an individual would prefer not to say, or if an individual does not consider themselves to have a gender identity. Producers should also consider the design of the question, including the benefits and costs of using free-text boxes, as well as whether specific questions should be age restricted or be voluntary. These approaches may differ depending on different respondent groups or collection modes.

Producers should consider the mode of data collection, including when there is a change of mode, or when collecting data on sensitive topics. It may be that a respondent would answer differently to a question depending on whether the data were collected by an interviewer face to face or if the questions were answered online. Responses may also vary depending on whether the data are self-reported or completed by another individual, such as a household representative.

Producers who are collecting data through statistical surveys may have more involvement and control in making decisions on the inclusion of questions and question testing and design. If the data are collected through administrative systems, it may be more complex to add, adapt or improve questions, and there may be significant costs involved. This is

because these data are usually collected for operational purposes and statistical reporting data requirements may be secondary. If administrative data collections cannot be changed, producers should assess whether this affects the suitability or quality of data for statistical reporting purposes.

Where data are collected through an administrative system, producers should establish and maintain a constructive relationship with the data suppliers. They should seek to understand how the data are collected, discuss any data requirements and understand any quality assurance or adjustments that have been carried out on the data, all within the particular operational context. Producers should also have a mechanism for two-way communication on any issues such as data errors or other quality concerns and keep informed of any changes to data collections.

Producers should consider how data about sex and gender identity will be recorded in administrative systems and how this may affect the responses received as well as the acceptability of questions, for example, whether the data are self-reported or completed by another individual such as an operational official. Producers should consider the impacts of this on the resultant data.

## Questions for statistics producers to consider

### Are you using survey data to produce your statistics?

Producers who source their data from surveys are likely to have more control over how information on sex or gender identity is collected, including a greater ability to add or update questions.

Ask yourself:

- Have you consulted your Head of Profession for Statistics, Chief Statistician or Lead Official for advice on what guidance on data about sex and gender identity may be relevant for your statistics?
- If you are adding a new question or updating an existing question on your survey, have you undertaken appropriate question testing?
- Have you considered how the mode of collection, for example, face to face compared to online, may affect the quality of the data collected?
- Have you provided guidance to respondents or interviewers about how to answer relevant questions and explained the definitions that should be used?

### Are you using administrative data to produce your statistics?

Producers who source their data from administrative systems are less likely to be directly involved in how this information is collected. This can make it more difficult for producers to understand how the information is collected and to introduce changes or improvements. For more detail on working with this type of data, we recommend reviewing our guidance on [Quality Assurance of Administrative Data \(QAAD\)](#) and the Government Analysis Function's [Quality of administrative data in statistics](#) guidance.

Ask yourself:

- Do you understand how information about sex or gender identity is recorded in your administrative data sources?

- Are you aware of existing operational guidance or metadata on how this information should be recorded?
- Have you identified any potential limitations of how the data are recorded?
- Are you working with relevant parties, where possible, to ensure that the data collection meets your needs?

# How to report and considerations for onward use of the data

## OSR's expectations

As with data collection, users of statistics benefit when producers are clear about what is reported. Data about sex and gender identity should be explained and defined for the purpose of a particular set of statistics, and terms, including 'gender', should not be used interchangeably or as a substitute for each other.

Producers must clearly explain the sources of data about sex or gender identity that they are using and how these data are collected. This applies to both survey and administrative data sources. Where producers do not have a clear understanding of how this information has been collected, this should be communicated transparently to users. Producers should also provide information about how responses such as 'prefer not to say' are treated in data processing.

The quality and any limitations of the source data should be shared with users. For surveys, this may include information on response rates, sample size, response bias and mode effect. For administrative data sources, this may include information on the potential sources of bias and error in administrative systems and how complete the data are.

If changes have been made to the data collection, including the questions asked or the nature of the guidance provided, this should be explained clearly to users, including information on any potential impact on comparability over time.

Many surveys are weighted to increase the representativeness of the data as well as to mitigate the impact of non-response bias. Producers who are applying weights to their survey data should ensure that the concepts within the weighting dataset and the survey data are aligned as far as possible. Producers should clearly explain the nature and the source of the data and the methodology used.

Producers should carefully consider how they assure themselves of the quality of the statistics, particularly where questions are voluntary or there may be limited other sources of data to compare or triangulate with. Producers should clearly explain the extent and nature of any uncertainty in the statistics.

The strengths and limitations of the published statistics should be considered in relation to different uses and be clearly explained. This may be particularly relevant when considering data at small sub-group levels, for example for certain gender identities. Producers should support users in understanding the data and the conclusions that can be drawn.

Producers should apply appropriate statistical disclosure control methods to protect individuals' data. This is especially important when producers are reporting on small populations. If producers are reporting on gender identity, they may find that some of the response options are too small to report on and the data may need to be combined or suppressed.



## Onward use

We strongly champion the onward use of data including through data sharing, data linkage and the reuse of data where feasible. We consider that these activities greatly increase the value of the statistics and data.

When providing data for onward use, such as through Trusted Research Environments, producers should ensure that they provide clear information on how data about sex or gender identity have been collected and processed. This information will help analysts and researchers understand the approach taken for each data collection.

When using combined datasets that may allow individuals to be classified as transgender, producers should be mindful of ethical considerations, seeking advice from their Head of Profession for Statistics, Chief Statistician or Lead Official and the expertise of the UK Statistics Authority's Centre for Applied Data Ethics if necessary.

## Questions for statistics producers to consider

### How well are you communicating information to users?

Producers should ensure that statistics, data and explanatory materials are presented in a clear, unambiguous way that supports their intended use. The strengths and limitations of the statistics and data should be clearly explained to users, and users should be made aware of the nature and extent of any changes to the data collection.

Ask yourself:

- Have you been clear about the definitions or terms that you have used?
- Have you communicated any changes to the data collection to users?
- Have you included information on how these data are collected and any associated limitations?

### Have you included key information about your statistical methods?

Producers should ensure that they are transparent about their methods and provide key information to support users when interpreting the statistics or conducting their own analysis.

Ask yourself:

- Have you explained to users the extent to which your approach to collecting data about sex or gender identity is coherent with related statistics?
- If you are applying weights to your data, have you assessed how closely the source data align to the concepts within your survey? Have you also clearly explained your weighting methodology as well as any associated limitations?
- Have you applied appropriate statistical disclosure control methods?

# Understanding user needs

## OSR's expectations

The Code states that statistics should meet their intended uses and should inform public debate. To achieve this, producers must seek to understand their users and the questions that users want to be able to answer with their statistics.

Statistics producers should seek a wide range of user views and consider the different uses of their data. There are many ways to do this, including actively approaching users for feedback on proposals; publishing work plans; undertaking consultations; conducting research exercises; monitoring public debate; and considering advancements or developments in other areas of data and statistics provision. Producers should consider if there are opportunities to work collaboratively when developing user engagement strategies to maximise benefits across the wider statistical system and reduce burden on users and stakeholders. Our [regulatory guidance on user engagement](#), as well as the Government Analysis Function [guidance](#) on user engagement, provides helpful practical advice on understanding user needs.

The needs of people using data and statistics can be extensive and varied. This can make it difficult for producers to meet the needs of all users with the same output. We also recognise that the collection and reporting of data about sex and gender identity is a sensitive topic which presents further challenges for producers. It is important that producers understand the current wider societal issues and debates about sex and gender identity. When necessary, producers should access a range of expertise across the statistical system on these issues to support their understanding and decision making.

Statistics should be regularly reviewed, involving users and other stakeholders, to prioritise development plans. Decisions about whether to continue, adaptor discontinue statistics about sex or gender identity should be made with stakeholders.

When an evolving or new user need is identified, producers should consider whether the data currently collected can and should be enhanced or if new data requirements could be met through other opportunities such as data linkage. Producers should work collaboratively, where appropriate, when reviewing and developing their statistics, including across countries or topic areas. Producers should consider the impact of any potential changes and whether they would result in the needs of some users no longer being met. Information on decision making relating to how users' needs can or cannot be met should be fed back.

## Questions for statistics producers to consider

### Do you have a good understanding of your user needs?

It is important that producers have a good understanding of what they are collecting and whether this meets the needs of their users. Often, data about sex are a long-standing part of the collection and the reasons for how or why these data are being collected and used may not have been considered for some time.

Ask yourself:

- Have you engaged with a wide range of users to understand their needs?
- Has your user engagement been balanced across different viewpoints where possible?
- Are there improvements or changes that you could make to your collection of data about sex or gender identity to better meet the needs of your users?
- If you have identified a user need but are unable to meet it, have you clearly explained the reason for this to users?

# Case studies

To support our guidance material would like to include case studies which exemplify certain elements of the guidance.

We encourage statistics producers to share their experiences more widely. We recognise many producers are navigating and seeking improvements to their data collections about sex and gender identity. Sharing collective learning, alongside our overarching guidelines on Code compliance as it relates to this topic, helps support systemic improvements to the production of data and statistics that serve the public.

If you would like to share your own experience as a producer and add a case study to this guidance document, please contact [regulation@statistics.gov.uk](mailto:regulation@statistics.gov.uk).

# What's changed in OSR's guidance?

This guidance updates and supersedes the OSR previous draft guidance on collecting and reporting data about sex in official statistics, published in 2021.

This guidance was developed based on feedback on our pre-existing guidance, as well as our regulatory work on these issues. We sought feedback on a draft version from official statistics producers across a wide range of government departments and the Devolved Administrations. We also met with a number of external stakeholders to gather wider views. We would like to express our gratitude to all the individuals who engaged with us, this process was instrumental in the development of this guidance.

Any future updates to this guidance will be recorded here.

1. First publication - 29 February 2024

2. Current version - 17 December 2024 update

- Amended the terminology section for gender identity to read as follows - .  
*Gender identity is not constrained to the binary male/man and female/woman categories and includes others such as non-binary and gender fluid.*

Previous text was, *Gender identity is not constrained to the binary male and female categories and includes others such as non-binary and gender fluid.*

- Moved the section “OSR programme of work on sex and gender identity” to earlier in the document. Updated the text in this section to reflect the publication of OSR's final report on the Gender identity review of the 2021 Census in England and Wales.
- Updated the section “Questions used to collect data about sex and gender identity” to now also reference new GSS Harmonisation publications. Removed the reference to the Census 2021 questions and retained the reference to the Scottish Government Chief Statisticians guidance.
- Whole document reviewed for readability and grammar (copy edit review) and minor edits made where these benefit the reader.