Understanding Equality Data Collection in the Scottish Public Sector

EQUALITY AND WELFARE

social research

Understanding equality data collection in the Scottish public sector

Final Report

March 2021

Jennifer Waterton, Alison Platts, Dawn Griesbach

Jennifer Waterton Consultancy

Table of contents

Acknowledgements

List of acronyms

The following acronyms are used in this report.

GDPR	The European Union's General Data Protection Regulation 2016/679 is a regulation in EU law on data protection and privacy in the European Union and the European Economic Area. It stipulates a variety of requirements around how and why personal data can be processed. It came into force on 25 May 2018.
DPA 2018	The Data Protection Act 2018 (DPA 2018) is a United Kingdom Act of Parliament which is the UK implementation of the EU's GDPR legislation, codifying its requirements into UK law. It came into force on 23 May 2018.
PSED	Public Sector Equality Duty - a duty on public authorities to consider or think about how their policies or decisions affect people who are protected under the Equality Act.
EMF	Equality Monitoring Form – a self-contained set of equality questions that can be used in relation to a number of different processes or services.

Executive Summary

1.   The Scottish Government is committed to tackling inequalities in Scotland. This has been a central element in its broad policy approach over recent years, as seen in the National Performance Framework, and underpinned by the Equality Act 2010.[1,2] In 2020, as a result of the Covid-19 pandemic – and concerns about its impact on different equality groups and the socio-economically disadvantaged – equality issues have come into sharp focus.

2.   In order for the Scottish Government's ambition for the advancement of equality to be realised, it is vital that robust and comprehensive data on the characteristics of the people of Scotland (including equality and socio-economic disadvantage data) are collected. The availability and use of such data allows policies and services to be developed, delivered and adjusted in ways which align with the ambition to reduce inequalities.

3.   In September 2020, the Scottish Government commissioned research to support its efforts to enhance the availability and quality of equality and socio-economic disadvantage data collected by public sector organisations in Scotland.

Aims of the research

4.   The research had two aims. The first aim was to identify, describe and explore (i) what equality and socio-economic disadvantage data are collected by public sector organisations,[3] (ii) how equality data[4] are collected and stored, (iii) why or for what purpose equality data are collected and whether that purpose is fulfilled, and (iv) the barriers to equality data collection and to its use for the outlined purpose. The second aim was to use the findings of the research to offer insights into what works best in terms of collecting, utilising and safeguarding robust equality data and to highlight major barriers that could prevent collection or use of these data. The research also investigated views on whether the collection of equality data should be 'more extensively mandated under legislation' as has been suggested by some stakeholders in this field.

Research methods and approach

5.   The research involved exploring the collection of equality data within a sample of Scottish public sector organisations and networks, gathering descriptive information about a range of identified data collections, and exploring the views and experiences of individuals working with selected data collections in a variety of capacities.[5]

6.   The research focused on organisations (and policy areas) with particular relevance to addressing 'structural inequality'[6] and on 'public facing' organisations (i.e. service providers). The achieved sample comprised 27 organisations working across a range of

---

[1] Scottish Government, National Performance Framework.
[2] Equality Act 2010.
[3] For the purpose of this research the term 'data collection' is restricted to quantitative data collected through surveys, equality monitoring forms or administrative processes.
[4] From here onwards, the phrase 'equality data' is used to mean 'equality and socio-economic disadvantage data'.
[5] Note that the Data Protection Act 2018 provides the current legal framework for the collecting and processing of personal data in the UK. This Act codifies into UK law the European Union's General Data Protection Legislation.
[6] Specifically, Health and social care, Justice, Education and skills, and Employment, labour and welfare.

policy areas (e.g. health and social care, justice, education and skills, etc.). In all, 55 datasets were included in the research.

7.   The study involved: (i) a desk-based review of organisations' websites and publications, (ii) the collection of descriptive information about each selected data collection, (iii) qualitative interviews to explore issues related to the collection of equality data, and (iv) discussions with public sector equality networks.

What equality data are collected?

8.   The research found that data on age and sex were collected in almost all of the data collections included in the study. Race and disability were collected in a large majority of cases. Religion and belief, and sexual orientation were collected in more than a third of cases. Marriage and civil partnership and gender reassignment were collected in around a quarter of cases, and pregnancy and maternity was collected in a small minority of cases.

9.   One collection contained information about all nine (9) 'protected characteristics'[7] and 14 further collections contained information about seven (7) or more characteristics. By contrast, seven (7) collections contained information about two (2) or fewer characteristics.

10.   As far as socio-economic disadvantage data were concerned, three-quarters of the data collections contained postcode information (which can, at least in theory, be linked to the Scottish Index of Multiple Deprivation[8]), around a quarter collected data on household incomes, and around a quarter collected data on other indicators of socio-economic disadvantage (e.g. property type, occupation, employment status, etc.).

11.   In some cases, data on other equality-related characteristics were also collected. Care experience / care leaver and caring responsibilities were the most frequently reported additional characteristics collected; these were often collected in relation to education and other services with a primary focus on children and young people.

12.   The question wording, accompanying definitions, and response categories were diverse, detailed and often complex. In general, there was a lack of standardisation in the way items are collected, in terms of: (i) the wording of / terminology used in the question, (ii) the definitions supplied to support the question, and (iii) the range of response categories offered. However, there were also a lot of similarities in the questions asked, and the differences were often detailed in nature.

13.   Participants in the study often said that the Census 2011 questions had provided a reference point for developing the wording, definitions and response categories for the equality questions. There was also widespread use of a 'prefer not to say' response category (especially in relation to religion, sexual orientation, and gender reassignment).

---

[7] The nine (9) protected characteristics defined by the Equality Act 2010 cover: age, religion and belief, race, disability, sex, sexual orientation, pregnancy and maternity, marriage and civil partnership, and gender reassignment.
[8] The Scottish Index of Multiple Deprivation (SIMD) classifies small geographical areas (called 'data zones') based on information across seven domains: income, employment, education, health, access to services, crime and housing.

Characteristics of equality data collections, and their purpose and uses

14. The (55) data collections explored in the research were highly diverse in terms of their content, operational context, and coverage. There was also variation in relation to the amount of 'infrastructure' (i.e. expertise, resources, and other supporting arrangements, both 'in-house' and external to the organisation) available to support their collection and use.

15. Around a third of the data collections were described as 'surveys' with the remainder described as 'administrative data'. Around a quarter were designated as 'National Statistics' (which means they meet the standards set out in the UK Statistics Authority Code of Practice).[9] Approximately one-half were Scotland-wide, with a further one-fifth based in a single local authority. Two-thirds (including most of the administrative data collections) were described as 'ongoing' data collections.

16. The study found that Scottish public sector bodies collect and use equality data for a wide variety of purposes. In broad terms these purposes relate to: (i) fulfilling statutory requirements, (ii) developing, monitoring, evaluating and improving (national and local) policy, and (iii) developing, delivering, monitoring, evaluating and improving services.

17. The study found little evidence of data being collected, but not used. However, there was a range of cases where interviewees said the organisation could make more use – or better use – of the equality data they collected.

How are equality data collected and processed?

18. The methods used by public sector organisations to collect equality data included online, telephone, face-to-face, and paper-based methods – sometimes used in combination. There was also variation in (i) whether the information was provided to the organisation by the customer, or gathered and recorded by a 'third party' (i.e. a partner or external organisation, or individual(s) contracted to provide services), and (ii) the extent to which the collection of equality data was described as 'voluntary' or 'mandatory / required'.

19. Research participants also described the use of equality monitoring forms (EMFs).[10] This took three main forms: firstly, EMFs formed the basis of the data collection; secondly, a separate EMF accompanied the 'main' data collection (note that in some cases these were returned separately); and thirdly, the EMF was integrated within the 'main' data collection.

20. The processing of equality data by public sector organisations (e.g. data input, quality control, data cleaning, data revisions and storage) is highly dependent on the infrastructure and resources which are available to support the particular data collection. Compliance with the Data Protection Act 2018 (DPA 2018) is the basic legal requirement, and the organisations participating in this study were very conscious of this.

---

[9] Office for Statistics Regulation, National Statistics.
[10] An equality monitoring form is a self-contained set of equality questions that can be used in relation to a number of different processes or services.

21.   The research heard about very varied processes. On the one hand, some data collections were gathered on paper (and in most, but not all cases, subsequently input to an electronic system). In other cases, there were very detailed arrangements for quality control and feedback (involving several iterations between respondents and the bodies collecting and reviewing the information), and collections were integrated into extensive IT systems and platforms, with complex and carefully developed data linkage arrangements in place.

Barriers and challenges in collecting equality data

22.   The main barriers and challenges identified related to: (i) the personal and sensitive nature of equality data, (ii) practical, operational and / or methodological issues, (iii) difficulties relating to definitions and terminology, and (iv) shortcomings in the organisational culture, capacity and / or capability.

23.   As far as the **personal and sensitive nature of equality data** is concerned, research participants explained that people can be reluctant to provide information about their equality characteristics, and may be concerned that this information might be used to discriminate against them. Members of staff can also feel uncomfortable asking people about their protected characteristics. There was also a view that individuals are less likely to disclose personal and sensitive information if they did not trust (i) the organisation asking for the information or (ii) the rationale given for collecting it. More generally, participants talked in terms of 'tensions' in relation to the collection of these data, the 'balances which needed to be struck' and the importance that the collection of equality data should be 'proportionate'.

24.   In terms of **practical, operational or methodological challenges**, participants explained that the social and emotional context and /or the physical environment within which data collection takes place is not always conducive to the collection of high-quality data. For example, data collection can take place during a medical emergency or a police incident; outdoors in poor weather; in busy reception areas or other public areas; etc. Other challenges included: the lack of time available to collect data; the requirement to collect data by observation rather than by direct questioning; 'third party' data collection; and the multiple interpretations of some of the terminology used in equality data collection. This could be a particular issue for organisations providing 'frontline' services and dealing with the public in difficult or highly charged situations.

25.   The third challenge – relating to **definitions and terminology** – was linked to a lack of standardised approaches in (i) developing questions, (ii) defining response categories, (iii) using terminology, and (iv) the definitions which underpin the questions. This was seen as raising issues of inconsistency across the public sector and between partner organisations. These challenges were raised most often in relation to data on sex and gender. Participants also said it could be difficult to keep abreast of changes in best practice in collecting equality data.

26.   As far as **shortcomings in organisational culture** were concerned, a range of specific difficulties were described in which equality issues and the collection of equality data (i) were not prioritised or fully understood by senior management, (ii) were not

embraced beyond the need to meet legal requirements, or (iii) had been side-lined by other organisational pressures.

27.   In terms of the **shortcomings in capacity and / or capability**, some participants in this study highlighted (i) the lack of a 'joined up' approach to collecting and using equality data across the organisation, (ii) inadequate expertise for designing research tools, (iii) insufficient resources for data input, quality control and data processing, and (iv) a lack of analytical capacity to maximise the use / benefit of the data which were collected.

Enablers and facilitators in equality data collection

28.   Research participants described factors that enabled or facilitated (or could enable or facilitate) the collection of high-quality equality data. These covered: (i) mainstreaming equality, (ii) the importance of being clear about the purpose of collecting data, (iii) the importance of training (for staff and third-party data handlers), (iv) using data collection methods that work, (v) building data management and analytical capacity, and (vi) improving guidance and developing networks for support.

29.   Some participants highlighted their **organisation's efforts to 'mainstream' equality**. There was no standard approach to this, but suggestions for progressing this agenda focused on organisational structures, training, and leadership. Participants also emphasised the importance of taking steps to change organisational culture and behaviour. The requirement for public sector bodies to produce regular equality mainstreaming reports helped reinforce an emphasis on the development of equality outcomes.

30.   In the main, organisations that were collecting high-quality equality data with good response rates had invested time and effort in developing **clear communication with their customers**. For example, they provided a short explanation, in plain English, about why they were asking for this information, what the information would be used for, and how the respondent's privacy in relation to the data would be protected.

31.   Organisations that had high-quality equality data collections had also invested time in **training staff and third-party data handlers** to collect the data. This included the development of guidance, training materials, and 'scripts' that staff could refer to when speaking to customers.

32.   In terms of **data collection methods,** organisations that collected equality data directly from their customers through an online system generally reported higher quality, more complete data. The advantages of online data collection included: (i) the facility for individuals to update their own information, (ii) the (in general) higher quality of self-reported data (compared with data collected through an intermediary or through observation), (iii) greater efficiency and cost-effectiveness, and (iv) the greater ability to build quality control procedures into data collection processes.

33.   Where organisations had successfully **built capacity for data management, analysis and reporting**, this often involved the development of new IT systems and platforms incorporating better online data collection facilities; improved data cleaning; enhanced analytical capability; and more comprehensive reporting. These enhancements

could create a 'virtuous circle' whereby the demonstration of potential benefits to the organisation of collecting and using the data could lead to improved collection and use.

34.   Organisations with high-quality data collections – and those who wanted to improve their data collections – generally supported the idea of **improved guidance and the development of networks**. Suggestions about how this could be done included: (i) developing (better) standardisation and coordination of equality data collection – both within and across organisations, (ii) establishing and participating in (both internal and external) 'equality fora', and (iii) establishing a (nationally coordinated) 'go-to' place for guidance and support.

Views on making equality data collection mandatory

35.   There was a range of views in relation to the question of whether equality data collection should be made mandatory. Whilst some participants were open to this suggestion and / or welcomed a move in this direction, others were more sceptical and had reservations and concerns of various types. Whether participants were in favour or not, they identified resource implications, if this option were to be pursued.

36.   Those with **positive or neutral views** regarding mandatory equality data collection highlighted three main points. First, in highly developed administrative data systems there are detailed processes (including consultation with stakeholders) in place to determine what equality data are required. In such systems the quality of the data for mandatory fields tends to be very good. Second, in some specific cases (e.g. eligibility for financial or other support in relation to training or benefits) the data are already required (i.e. 'mandatory') for operational reasons. Finally, some organisations were keen to collect high-quality information about their 'customers' and so authorisation to collect this (if it were to be made mandatory by Scottish Government) would be welcomed / accommodated.

37.   Participants who expressed **concerns and reservations** about mandatory equality data collection highlighted a range of issues including that: (i) an 'inflexible' or 'blanket' approach was not helpful; these kinds of decisions should be made locally by those who understood the environment within which data collection takes place and the requirements for the organisation to collect equality data, (ii) such a requirement would (or could) be disproportionate or inappropriate or the benefit of it for an organisation was not clear, and (iii) there could be a reluctance by organisations to collect equality data if previously undertaken analysis had revealed no policy impacts in relation to specific characteristics. It was also argued that any mandating of approach would present difficulties and / or would be inappropriate for services that deal with the public in more challenging situations.

Reflections on 'what works best'

38.   Drawing on the research findings, four key themes were identified in relation to 'what works best'.

39.   First, it was clear that the **mainstreaming equality** agenda – including the requirement for regular reporting – has been an important driver in promoting a focus on, and improving, equality data collection. Organisations had, for example, responded positively to this agenda by developing a supportive organisational culture and proactive

leadership, investing in infrastructure, and ensuring a joined up, cohesive approach to equality data collection. In addition, it was thought that there was an important role for the Scottish Government to play in providing more guidance and improving the consistency of the data collected.

40.   Second, there was compelling evidence of the importance of ensuring that all those involved in developing equality data systems were **clear** about why questions are asked, and how the information gathered will be used. This clarity is already a requirement of the DPA 2018. However, if equality data collection is to be improved, the need for clarity goes far beyond a (narrow) legal requirement. This clarity can build trust and improve disclosure. It is often prompted by – and developed as a consequence of – 'normalising' and 'mainstreaming' the equality agenda within organisations.

41.   Third, this study found that public sector equality data collections were (not surprisingly) highly diverse in terms of subject matter. But they were also **diverse in terms of their operational contexts**. This variation was seen in terms of (i) the physical locations and emotional and social contexts in which data were collected, (ii) the individual(s) carrying out the data collection, and (iii) (partly as a consequence of (i) and (ii) above) the methodological approaches to collecting the data. This means that any approach to equality data collection must be carefully tailored to the situation at hand. The task of improving equality data collection cannot take a 'one size fits all' approach.

42.   Finally, equality data collection is a **highly complex undertaking**. Larger organisations have a variety of internal and external networks and sources of advice they can access on an ongoing basis. By contrast, other organisations, including, but not exclusively, smaller organisations, felt there was a lack of a **'go-to' place for advice and expertise on equality data collection**. This was notwithstanding that many of these organisations belonged to some kind of 'Equality Network' which they found helpful. Organisations with less equality data collection 'infrastructure' at their disposal had many questions about the 'nuts and bolts' of how best to undertake this work, and how to keep up with the (rapid) developments in this field.

1 Introduction

1.1   In September 2020, the Scottish Government commissioned research to improve understanding of the collection of equality and socio-economic disadvantage data by Scottish public sector bodies.

1.2   The research was intended to (i) improve the Scottish Government's understanding of the collection and use of these data in policy making and service design (and delivery) across public services, (ii) explore the barriers to data collection, and (iii) provide insights into what works best in terms of collecting, utilising and safeguarding robust equality and socio-economic disadvantage data.

1.3   This report sets out:

- The background to the project and the aims of the research (Section 2)
- The research methods and approach adopted in undertaking the research (Section 3)
- The findings of the research (Sections 4 to 9), which covered:
  - What equality and socio-economic disadvantage data are collected by public sector bodies (Section 4)
  - The characteristics of equality and socio-economic disadvantage data collections, their purposes and uses (Section 5)
  - How equality and socio-economic disadvantage data are collected and processed (Section 6)
  - Views and experiences in relation to the barriers and challenges in undertaking the collection of equality and socio-economic disadvantage data (Section 7)
  - Views and experiences in relation to the enablers and facilitators in undertaking the collection of equality and socio-economic disadvantage data (Section 8)
  - Views on making the collection of equality and socio-economic disadvantage data mandatory (Section 9)
- Reflections in relation to 'what works best' in terms of collecting, utilising and safeguarding robust equality and socio-economic disadvantage data (Section 10).

1.4   Information on participating organisations, the data collections included, and the research data collection tools are set out in Annexes 1 to 5. Annex 6 provides details about the capture of individual equality data items.

2 Background

2.1   This section describes (i) the policy context for the research, (ii) the legal framework within which the collection and processing of personal data takes place, and (iii) the aims which the research is intended to address.

Policy context

2.2   The Equality Act 2010 came into force in England, Wales and Scotland on 1 October 2010.[11] The Act brought together over 100 separate pieces of legislation (including the Sex Discrimination Act 1975, the Race Relations Act 1976, and the Disability Discrimination Act 1995) and provides a legal framework for protecting the rights of individuals and advancing equality of opportunity for all.

2.3   The 2010 Act explains that having due regard for advancing equality involves:

- Removing or minimising disadvantages suffered by people due to their 'protected characteristics'[12]

- Taking steps to meet the needs of people from protected groups where these are different from the needs of other people

- Encouraging people from protected groups to participate in public life or in other activities where their participation is disproportionately low.

2.4   The (Scottish) Public Sector Equality Duty is set out in the Equality Act 2010 and the (Specific Duties) (Scotland) Regulations 2012.[13] These regulations require (most) public sector bodies in Scotland to (i) eliminate discrimination, harassment and victimisation, (ii) advance equality of opportunity between different groups, and (iii) foster good relations with regard to the nine (9) protected characteristics of age, religion and belief, race, disability, sex, sexual orientation, pregnancy and maternity, marriage and civil partnership, and gender reassignment.

2.5   In addition, the Fairer Scotland Duty (Part 1 of the Equality Act 2010) places a legal responsibility on particular public bodies in Scotland to actively consider how they can reduce inequalities of outcome caused by socio-economic disadvantage when making strategic decisions. This came into force in Scotland in April 2018.[14]

2.6   An ambition to advance equality lies at the heart of the Scottish Government's vision for Scotland. This commitment is articulated through Scotland's National Performance Framework and specifically through the Framework's:

- **Purpose** – 'to reduce inequalities and give equal importance to economic, environmental and social progress'

---

[11] Equality Act 2010.
[12] The Act provides a range of protections for nine 'protected characteristics': age, religion and belief, race, disability, sex, sexual orientation, pregnancy and maternity, marriage and civil partnership, and gender reassignment.
[13] The Equality Act 2010 (Specific Duties) (Scotland) Regulations 2012.
[14] Scottish Government (2018) Fairer Scotland Duty: interim guidance for public bodies.

- **Values** – 'to treat all our people with kindness, dignity and compassion' and

- **Outcomes** – which 'reflect the aspirations and values of the people of Scotland, are aligned with the United Nations Sustainable Development Goals, and help to track progress in reducing inequality'.[15]

2.7   In order for the Scottish Government's ambition for the advancement of equality to be realised, it is vital that robust and comprehensive data about the social and economic characteristics of the people of Scotland (including equality and socio-economic disadvantage data) are collected. The availability and use of such data allows policies and services to be developed, delivered and adjusted in ways which align with the ambition to reduce inequalities, whilst its lack or absence will mean that it is not possible to recognise and address shortcomings, or track progress towards the vision for Scotland which has been identified.

2.8   As part of their efforts to address the issues around the collection and availability of equality data, the Scottish Government developed its Equality Evidence Finder.[16] This resource first appeared in 2012 and is updated twice a year. The Equality Evidence Finder provides a range of equality data across many policy areas, drawing on an extensive array of sources. The Scottish Government also published an *Equality Evidence Strategy* in 2017 which sets out its approach to improving the equality data which are available.[17]

2.9   Equality issues came into sharp focus for the Scottish Government in 2020 as a result of the Covid-19 pandemic. In June 2020, the Scottish Government published an analysis of the health and social impact of Covid-19 in relation to protected characteristics and socio-economic disadvantage.[18] Throughout the pandemic, concerns about the impacts on older people, disabled people and those with underlying health conditions, people from black and minority ethnic communities, and socio-economically disadvantaged people have featured heavily in public discourse.

Collecting and processing personal data: the current legal framework

2.10 The Data Protection Act 2018 (DPA 2018) is a United Kingdom Act of Parliament which updates data protection laws in the UK.[19] It is a national law which codifies into UK law the European Union's General Data Protection Regulation (GDPR).[20]

2.11 The DPA 2018 (and the GDPR) requires organisations that collect and process personal data to follow strict rules called 'data protection principles'. They must make sure the information is:

- Used fairly, lawfully and transparently

- Used for specified, explicit purposes

- Used in a way that is adequate, relevant and limited to only what is necessary

---

[15] Scottish Government, National Performance Framework.
[16] Scottish Government, Equality Evidence Finder.
[17] Scottish Government (2017) Scotland's equality evidence strategy 2017-2021, section 6: Scope of the Strategy.
[18] Scottish Government (2020) *Coronavirus (COVID-19): Health and social impact assessment.* June.
[19] Data Protection Act 2018.
[20] General Data Protection Regulation (GDPR).

- Accurate and, where necessary, kept up to date

- Kept for no longer than is necessary

- Handled in a way that ensures appropriate security, including protection against unlawful or unauthorised processing, access, loss, destruction or damage.

2.12   Thus, the DPA 2018 (and the GDPR) plays a vital role in relation to (i) what data are collected, (ii) how consent for data collection (and follow up) is obtained, (iii) how data are processed and stored, (iv) whether a Data Protection Impact Assessment[21] (DPIA) is required, and (v) how data are used and reported.[22]

Aims and objectives

2.13 The research described here is one element of a broad programme of work being pursued by the Scottish Government to enhance the collection, use, and reporting of equality and socio-economic disadvantage data in order to improve policy making, service design and delivery.

2.14 This research has five (5) aims as follows:

- Aim 1: To identify, describe and explore what equality and socio-economic disadvantage data are collected[23]

- Aim 2: To identify, describe and explore how equality and socio-economic data are collected and stored

- Aim 3: To identify, describe and explore why or for what purpose data are collected and whether that purpose is fulfilled

- Aim 4: To identify, describe and explore the barriers to equality data collection and to its use for the outlined purpose

- Aim 5: To use findings from aims 1 to 4 to offer insights as to what works best in terms of collecting, utilising and safeguarding robust equality and socio-economic disadvantage data and to highlight major barriers that could prevent collection or use of these data.

2.15   The research also has an implicit aim to investigate views on whether the collection of equality and socio-economic disadvantage data should be 'more extensively mandated under legislation' as has been suggested by some stakeholders in this field.

---

[21] Also known as a Privacy Impact Assessment (PIA).
[22] The terms 'DPA 2018' and 'GDPR' have been used interchangeably by participants during this research. In what follows 'DPA 2018' is used as an umbrella term to cover both the DPA 2018 and the GDPR.
[23] For the purpose of this research the term 'data collection' is restricted to quantitative data collected through surveys, equality monitoring forms or administrative processes.

3 Research methods and approach

3.1   This section describes the research design and data collection methods used in the study. The section also covers (i) the limitations of the research, and (ii) the terminology adopted in the report. First, however, the approach taken to define the scope of the project is described.

Defining the scale and scope of the project

3.2   There are over 150 public sector bodies in Scotland involved in various capacities in the design and delivery of public services. In order to make the project manageable and to maximise its value, it was agreed that the research should:

- Be based on (a subset of) the data collections undertaken by around 30 public sector bodies

- Focus disproportionately on public sector bodies whose remit is associated with addressing structural inequality (specifically those involved in Health and Social Care; Education; Justice; and Employment, Labour and Welfare)

- Focus on data collections which are 'public facing' and which support policy or service development and delivery

- Exclude consideration of equality and socio-economic disadvantage data collected in relation to the workforce in the selected public sector bodies

- Exclude consideration of UK-wide data sources (HMRC, DWP, etc.).

Research design

3.3   The research involved exploring the collection of equality data within a sample of public sector organisations and networks, gathering descriptive information about a range of identified data collections, and exploring the views and experiences of individuals working with selected data collections in a variety of capacities.

3.4   The research design employed a three-stage stratified approach to sampling. Whilst this design did not provide formal statistical representativeness, it allowed coverage of a range of (i) types of organisation (non-departmental public bodies, health bodies, executive agencies, etc.), and (ii) types of data collection (administrative, survey). It also covered all (nine) protected characteristics as well as socio-economic disadvantage to some degree.

3.5   The three stages of the sampling approach are described in further detail below, followed by a discussion of the approach to 'substitutions' where the original selection was not viable.

Stage 1 – A sample of organisations

3.6   Approximately 170 Scottish public sector bodies were identified as 'in scope' for this project.[24] These bodies were organised into six groups (or 'strata'). Four of these groups reflected to specific topic or policy areas, namely: (i) Health and Social Care (ii)

---

[24] A list of Scottish public bodies is available from the Scottish Government's National Public Bodies Directory.

Employment, Labour and Welfare (iii) Education and Skills, and (iv) Justice. The fifth group was reserved for local authorities, whose work spans multiple policy areas. The final group contained all 'other' in-scope organisations.

3.7   Twenty-eight (28) organisations were initially selected for inclusion in the research as follows:

- 5 Health and Social Care organisations

- 3 Employment, Labour and Welfare organisations

- 6 Education and Skills organisations

- 5 Justice Organisations

- 5 Local Authorities (including a mix of large and small, and urban and rural)

- 4 'Other' Organisations.

Stage 2 – A sample of departments within selected local authorities / NHS Boards

3.8   A second-stage selection process was implemented in relation to the local authorities (LAs) and NHS Boards.[25] The original intention was to select **two** departments in each local authority and NHS Board – although it was not possible to achieve this in every case (see paragraph 3.11 below).

Stage 3 – A sample of data collections

3.9   At the final stage, relevant data collections in each of the selected organisations or departments were identified. Between one (1) and four (4) data collections from any particular organisation or department were included in the research. In cases where there was more than one relevant data collection, the selection was guided by (i) a wish to focus on services provided to individuals (rather than organisations), and (ii) a wish to explore datasets of different kinds (e.g. administrative datasets, surveys, 'one off' exercises, etc.).

Substitutions

3.10 If it was not possible to identify – or confirm the participation of – relevant individuals within the selected organisation within an agreed timeframe, a substitution was made based on the stratified design. (For example, if it was not possible to identify suitable participant(s) within a selected health and social care organisation then a substitute organisation from within the health and social care grouping was selected as a replacement.) Substitutions were also made in cases where, following initial research, the original selection was deemed to be 'out of scope'.[26]

---

[25] It was not possible within the resource constraints of the project to include all data collections within these large and complex organisations.

[26] Because, for example, the original selection did not undertake any relevant data collection or did not have sufficient capacity to engage with the exercise.

Achieved sample

3.11 The final achieved sample comprised 27 organisations, including five (5) local authorities and one (1) NHS Board.[27] As explained above (paragraph 3.8) there was an intention to select **two** departments in each of these organisations. However, this did not prove to be possible in every case. Overall, eight (8) departments (or service areas) were included (comprising seven (7) local authority departments and one (1) NHS Board department).[28]

3.12 Annex 1 lists the 27 organisations that participated in the research.

3.13 Annex 2 lists the 55 datasets which were included in the research.

Data collection

3.14 The data collection for the research comprised four main elements:

- Desk-based review activities including:
  - Exploring the websites of selected organisations
  - Reviewing documents and publications for relevant material about the organisation's approach to the collection and use of equality and socio-economic disadvantage data
  - Identifying relevant individuals / departments for follow up.

- Collection of descriptive information on selected data collections. For each data collection considered, this involved:
  - Describing the main features of the data collection (type of data collection, frequency of collection, population coverage, geographic coverage, publication arrangements, methods of data collection, etc.) (See Annex 3 – Template 1)
  - Confirming which – if any – protected characteristics and socio-economic disadvantage data are collected (See Annex 3 – Template 1)
  - Recording (i) the questions used to elicit information about protected characteristics and socio-economic disadvantage, and any relevant definitions, (ii) the categories offered for responses, and (iii) the response (or completion) rates for specific items (if available). (See Annex 4 – Template 2).

- Qualitative interviews with staff in selected public bodies which explored (see Annex 5):
  - Reasons for collecting (or not collecting) specific equality and socio-economic disadvantage data
  - How data are used, and whether data are used for the stated purposes
  - Approaches to publication

---

[27] Note that it was particularly challenging to engage with NHS Boards during this period, due to the Covid-19 pandemic.
[28] Thus, the final number of 'entities' included in the research was 29. This comprised 21 'unitary organisations' (i.e. organisations where no second stage sampling was undertaken) and 8 'departments'.

- Challenges encountered in collecting equality and socio-economic disadvantage data
- Issues related to quality and completeness of the data
- Improvement work and shared learning with regard to the collection of equality and socio-economic disadvantage data
- Future considerations, including the potential for greater legislative requirements.

- Discussion with the Non-Departmental Public Bodies (NDPB) Equality Forum, and the Scottish Council Equality Network (SCEN) about all aspects of the collection and use of equality and socio-economic data.[29]

Limitations of the research

3.15   As set out above, the research is based on (i) a purposive sample of 27 public sector organisations, and (ii) discussion with two public sector equality networks. The research participants were employed in a wide variety of roles and included senior managers, policy specialists, analysts, information managers, equality officers, statisticians, researchers and administrators. The issues discussed were complex and wide-ranging – and were often quite specific to the organisation concerned. Thus, the findings reported here should be regarded as exploratory and indicative, rather than definitive or comprehensive.

A note on terminology

3.16   From this point on in the report, the abbreviation 'equality data' is used to refer to data concerned with **both** protected characteristics **and** socio-economic disadvantage. Where only **one** of these elements (**either** protected characteristics **or** socio-economic disadvantage) is specifically being discussed, this is made clear in the text.

3.17   The term 'customer' is used to refer to the wide range of service users, clients, patients, grant recipients, training participants, applicants, students, pupils, members of the public, etc. who are asked by public sector organisations to provide information about their equality characteristics.

---

[29] Note that these discussions were not focussed on any specific data collection, but were of a more general nature.

4 What equality data are collected?

4.1   This section discusses the equality data that are collected. It is based on information derived from the examination and exploration of 55 separate data collections (see Annex 2), using the research tools set out in Annexes 3 to 5.[30]

4.2   It is important to note that participants in the research emphasised that when their organisation was considering what data to collect, this was done with reference to the Data Protection Act 2018. Thus, equality data collection was undertaken for 'explicit purposes' and limited to 'only what is necessary' for business purposes. Participants also stressed the importance of being clear from the outset how the equality data collected would be used.

> **Examples**
>
> It is not sufficient justification that something is 'interesting to know'. It has to be something which is required for the delivery of a service.
>
> We don't feel we should be asking for more – the information needs to be relevant, and needs to meet DPA requirements.

Collection of information about protected characteristics

4.3   Table 4.1 below shows the extent to which each of the (9) protected characteristics were collected in each of the 55 collections which this study considered. It can be seen that:

- Age and Sex were collected in almost every case

- Race and Disability were collected in a large majority of cases

- Religion and belief and Sexual orientation were collected in a substantial minority of cases (more than a third but less than a half); Marriage and civil partnership, and Gender reassignment were collected in around one-quarter of cases

- Pregnancy and maternity was collected in a small minority of cases (7).

Table 4.1 – Number of data collections containing each protected characteristic

Protected characteristic	n
Age	53
Religion and belief	22
Race	43
Disability	42
Sex	48
Sexual orientation	24
Pregnancy and maternity	7
Marriage and civil partnership	12
Gender reassignment	14

---

[30] In two data collections, the information gathered reflected changes due to be implemented in the coming months.

4.4   The following points should be noted with regard to the information in Table 4.1:

- 'Marriage and civil partnership' is covered by the Public Sector Equality Duty (PSED) **only** in relation to unlawful discrimination in employment. Although this study did not consider workforce / employment data (see paragraph 3.2 above), this protected characteristic was included in the current study for completeness.

- In addition to the numbers reported in Table 4.1 above, interviewees explained that, in relation to some data collections, information on a particular protected characteristic might be collected if it were relevant, and recorded in narrative or open text fields on case management systems or within files associated with an individual's case. However, the organisation's data collection protocol did not require this information to be systematically collected (and recorded) in all cases. (As a result, these cases are not included in the counts presented in Table 4.1.)

- In a few cases, public sector bodies sought information from organisations as well as individuals. In one case, for example, where both individuals and organisations were able to apply for grants, organisations were asked how many staff, board members, volunteers, etc. had a certain equality characteristic, with the options offered being the same as those offered to individual grant applicants. This point is returned to later in the report (see paragraph 7.18). Group level information was also sought in a few cases, for example, asking if anybody in a group of individuals had a disability.

4.5   Table 4.2 below presents the information gathered in this research in a slightly different way, by looking at the number of protected characteristics collected in each data collection. It can be seen that just one (1) collection contained information about all nine (9) protected characteristics. However, 15 collections contained information about seven (7) or more characteristics. By contrast, seven (7) collections contained information about two (2) or fewer characteristics.

Table 4.2 – Number of protected characteristics collected in each data collection

Number of protected characteristics	Number of data collections
0–2	7
3	13
4	7
5	8
6	5
7	10
8	4
9	1

Collection of information about socio-economic disadvantage

4.6   The indicators of socio-economic disadvantage considered in relation to this research were (i) area deprivation as measured by the Scottish Index of Multiple Deprivation (SIMD), and (ii) (any measure of) household income.

4.7   The SIMD classifies small geographical areas (called 'data zones') based on information across seven domains: income, employment, education, health, access to services, crime and housing. Thus, in order for SIMD analysis to be possible, postcode information is required to link an individual to the data zone in which they live.[31]

4.8   As far as the collection of these data are concerned it was found that:

- 41 of the collections included postcode information – although it was often clear that the postcode was **not** being used by the organisation as an indicator of socio-economic disadvantage (i.e. the postcode information was not being used analytically).
- 14 of the collections included information about household incomes.

4.9   Whilst it was not part of this study to explore indicators of socio-economic disadvantage other than SIMD category / postcode, or household income (see paragraph 4.6 above), it was found that in 15 of the data collections, **other information** relating to socio-economic disadvantage was collected.[32] This included information about household type; type of property; occupation (of the service user, or parental occupation in some cases); employment status; parental education; location of secondary school; and (non-income related) eligibility for assistance.

Collection of data about other 'equality' characteristics

4.10 In addition to the characteristics described above, the collections that were reviewed also collected information on a range of other characteristics which can be seen as 'equality-related characteristics' in a broad sense including:

- (Household) vulnerabilities
- Primary reason for contacting a service
- Communication requirements (including need for an interpreter / preferred language / language spoken in the home / British Sign Language requirement)
- Whether an individual is a Gaelic speaker
- Whether an individual is care experienced / or is a care leaver
- Whether a child is on the Child Protection Register
- Caring responsibilities

---

[31] Note that (i) collecting postcode information (ii) using it to generate a SIMD code and (iii) undertaking analysis based on SIMD are three separate issues. This project has identified whether postcode data were collected in each case; however the information about generating a SIMD code and undertaking SIMD-based analyses was addressed qualitatively (i.e. no counts were obtained for these aspects).

[32] Note, though, that in some cases, the information was collected as qualitative – rather than categorical – data.

- Whether a university / college student is estranged from their family
- Whether an individual is an armed services veteran.

4.11   Care experience / care leaver and caring responsibilities were the most frequently reported additional characteristics collected; these were particularly common in relation to education and other services with a primary focus on children and young people.

Equality characteristics – definitions and response categories

4.12   For data collections containing information about a specific protected characteristic or socio-economic disadvantage indicator, (i) the question which was used (including any definitional issues), and (ii) the response categories offered were ascertained.

4.13   The question wording, accompanying definitions, and response categories were diverse, detailed and often complex.

4.14   In general, there was a lack of standardisation in the way these items are collected, in terms of (i) the wording of / terminology used in the question, (ii) the definitions supplied to support the question, and (iii) the range of response categories offered. However, there were also a lot of similarities in the questions asked, and the differences were often detailed in nature.

4.15   Whilst some of the variation reflects the requirements of a specific data collection (for example, the definition of 'disability' employed and the response options offered may vary depending on whether the information is required to support decision making at individual level – e.g., in relation to a student's access to educational materials or a patient's requirement for transport to a health care setting), in many cases there was no obvious rationale for the differences.

4.16   It is also notable that there were variations **within** (as well as **between**) organisations in the way items were collected.

4.17   Two further general points should also be noted:

- Interviewees often said that the Census 2011 questions had provided a reference point for gathering equality data. However, even where questions were clearly based on census questions, there were often slight variations in wording or phrasing used or the response options offered.

- A 'prefer not to say' response option was often included, across the full range of equality characteristics. However, it was least likely to be offered in relation to age, and most likely to be offered in relation to religion, sexual orientation, and gender reassignment – with regard to these characteristics, there were very few examples of questions which did not offer this option. (Note that unless a question is set as mandatory, an individual can, of course, simply choose to not answer a question without a 'prefer not to say' option being offered.[33])

---

[33] Indeed, in some cases, a 'prefer not to say' option isn't included *because* a question isn't mandatory.

4.18   Annex 6 presents an overview of the questions, definitions, response categories – and variation on these – in relation to each individual protected characteristic.

Response rates and data 'completeness'

4.19   It was not always possible to get information about the completeness of equality data held by public sector bodies. Completion / responses rates often included 'prefer not to say' options and were also affected by whether a question was mandatory or voluntary. Organisations often reported good data completeness – 100%, or close to 100% – for questions or fields related to age, sex, and postcode. However, levels of completeness were much more varied, and occasionally much lower, for other characteristics – for example, a completion rate of 57% was reported for sexual orientation in one case, and in another a rate of 18% was reported for religion and belief.

4.20   Where organisations used separate equality monitoring forms, these typically had low response rates if their completion was voluntary. However, it was also noted by one interviewee that if an individual completed the form, they tended to complete all the questions.

5 Characteristics of equality data collections, and their purpose and uses

5.1   This section presents the findings of the research in relation to: (i) the characteristics of the data collections explored, (ii) why equality data are collected, and (iii) how they are (or are not) used. In order to set the context for this discussion, a brief account of the wide range and diversity of the data collections included in the research is first provided.

Diversity and range of data collections included in the research

5.2   The (55) data collections explored in the research are **highly diverse** in terms of their **content, operational context** and **coverage**. The data collections cover, for example, information related to:

- Transportation of patients to or between health care settings
- Visits by the general public to sites of historical and architectural interest
- Applications to and / or enrolments in a school, college, university or employment scheme
- The delivery of statutory interventions for children and adults
- The provision of grants, benefits or other types of funding
- Individuals in the criminal justice system
- Participation in sport by school aged children and young people
- Detention of individuals under the Mental Health Act
- Homeless applications.

5.3   These data collections can be thought of as being on a spectrum. At one end of this spectrum are data collections designated as 'National Statistics' (which means that they are fully compliant with the Code of Practice for Statistics as set out by the Office for Statistics Regulation and meet the highest standards of trustworthiness, quality and value). At the other end of the spectrum are data collections which have been developed within a single organisation for a specific purpose with no reference to any external adjudication or assessment. Between these two ends of the spectrum, there are a variety of data collections which are described as 'official statistics', that is, they are produced by crown bodies, those acting on behalf of crown bodies, or those specified in statutory orders.[34] The amount of 'infrastructure' and resource available to support the collection and subsequent analysis and reporting of these data collections (in terms of, for example, questionnaire design, methodological development, quality control, processing, IT support, analytical support, etc.) is different, depending on where on this spectrum a particular data collection lies.

---

[34] Office for Statistics Regulation, National Statistics.

Characteristics of the data collections explored

5.4 The data that were collected provide some basic descriptors which illustrate some of this diversity quantitatively (see Template 1, Annex 3). In particular:

- In relation to **type**:
    - 18 collections were described as 'surveys'
    - 37 collections were described as 'administrative' data collections
- In relation to **designation**:
    - 15 of the administrative data collections were designated as 'National Statistics'
- In relation to **geographic coverage:**
    - 27 collections were Scotland-wide
    - 10 collections covered (primarily) a single local authority
    - 8 collections related to some other geography (e.g. '10 local authority areas'; 'Forth Valley Region'; 'mainly Scotland')
    - 9 collections were not geographically defined / constrained (e.g. surveys of visitors to a national park; applicants to a college or university; experiences within a care home).
- In relation to **frequency**:
    - 36 collections were described as 'ongoing'[35]
    - 4 collections were (at least) annual
    - 6 collections were described as 'regular' (e.g. every 2 years; every 2-3 years; at the start of every term etc.)
    - 9 collections had some other cycle (intermittent; one-off; 'the first of two', etc.).

Why are equality data collected and what are they used for?

5.5 Scottish public sector bodies collect and use equality data for a wide variety of purposes. In broad terms these purposes relate to (i) delivering on statutory requirements, (ii) developing, monitoring, evaluating and improving (national and local) policy, and (iii) developing, delivering, monitoring, evaluating and improving services. Each of these categories is elaborated below with examples from the data collections explored.

5.6 Examples in relation to **delivering on statutory requirements** include:

---

[35] That is, new cases were added to the data collection in an ongoing way. Note that some of the 'ongoing' data collections could also be described as 'continuous' – where the record of any one individual was updated as more information became available. Examples of ongoing and continuous data collections included pupil records and social work case management information.

- Assisting with the production of Equality Outcomes and Mainstreaming reports which are required under the (Specific Duties) (Scotland) Regulations 2012. These reports are required on a regular (2–4 year) cycle[36]

- Producing or contributing to an official statistics publication as required by the Statistics and Registration Act 2007[37]

- Meeting reporting requirements set by (sometimes multiple) funding and sponsoring agencies and organisations

- Responding to Freedom of Information requests in relation to equality groups.

5.7   Organisations named a variety of specific pieces of legislation which provided an impetus for their collection of equality data. These included the Equality Act 2010, the Children and Young People (Scotland) Act 2014, the Mental Health (Care and Treatment) Scotland Act 2003, etc.

5.8   Examples in relation to **developing, monitoring, evaluating and improving (national and local) policy** include:

- Contributing to the identification of suitable / suitably targeted equality outcomes (e.g. educational attainment gaps, accessibility of healthcare services)

- Undertaking Equality Impact Assessments of proposed policy changes

- Measuring / assessing inequalities and developing plans and proposals to reduce / eliminate them

- Improving understanding of how different equality groups are / are not benefiting from funding / investment

- Promoting the collection and use of equality data to other relevant stakeholders

- Ensuring that policies are developed in a non-discriminatory way in relation to equality groups

- Providing high level performance management information or information for internal monitoring purposes.

5.9   Examples in relation to **developing, delivering, monitoring, evaluating and improving services** include:

- Determining eligibility for services or support (e.g. services may only be available to certain age groups, to people who are disabled, or to those with specified levels of socio-economic disadvantage)

- Improving understanding of the experiences of (services by) different users, and the differential impacts of experiences and services on equality groups (e.g. How is the access to and take-up of sporting opportunities affected by disability? How does the experience of emergency detention procedures vary by ethnicity? Are men more

---

[36] The Equality Act 2010 (Specific Duties) (Scotland) Regulations 2012.
[37] Statistics and Registration Service Act 2007.

23

likely than women to be awarded certain kinds of benefits? Are all equality groups able to access high quality outdoor experiences?)

- Understanding the profile of service users in relation to their equality characteristics

- Planning services in a way which ensures they are non-discriminatory in relation to equality groups

- Improving uptake of services by underrepresented groups

- Managing risk for those who need exceptional levels of care

- Tailoring training programmes and workshops so that they meet the needs of those attending

- Targeting literature at underrepresented groups

- Ensuring that all individuals / service users are supported in an appropriate way.

5.10   It is important to note that – particularly, but not only, in relation to the development of services (paragraph 5.9 above) – not all organisations would necessarily describe what they are doing as 'collecting equality data'. Some organisations were more likely to describe their activities as 'collecting data for service delivery, development and improvement' and the decisions about what to collect would be prompted by considerations about what was required for 'service reasons' rather than for 'equality reasons'. In other words, organisations do not always bring an equality perspective to the question of what data they collect (or should collect or might collect).

5.11   Other purposes and uses which have been identified include:

- To apply for awards or recognition in relation to the provision of policies and / or services to specific equality groups

- To support the work of / or report to external stakeholders (on a non-statutory basis)

- To inform the development of case management systems which include information about equality characteristics.

- To provide feedback to service providers or funders about performance in relation to equality groups.

5.12   Note that in terms of publication arrangements specifically, it was found that:

- 33 of the collections resulted in a publication about the data collection itself. However, not all the available equality characteristics were always analysed / reported.

- In the remaining 22 cases, there was no dedicated external publication produced. However, in some cases, information based on an analysis of the data collected was incorporated into other publications (e.g. corporate publications including statutory 'equality outcomes and mainstreaming' reports), or included within internal papers and reports, some of which may be available to the public (e.g. papers prepared for

consideration by boards or other strategic groups). However, an analysis of equality characteristics was not necessarily included in these reports.

Are there examples where equality data are collected but not used?

5.13   As discussed above, equality data were being used by organisations for a wide range of purposes, and there did not appear to be any widespread issues with data being collected but not used. There were, however, a few examples where equality data were collected but had not been used either at all or (in the interviewee's opinion) to their full extent. For example, one participant noted that end of project monitoring forms which recipients use to report on who has benefited from investment were not always fully scrutinised or analysed. A small number of organisations reported that 'there is probably information we are collecting that we are not analysing fully at the moment'.

6 How are equality data collected and processed?

6.1   This section briefly describes how equality data are collected and processed.

6.2   It is important to note that organisations often emphasised that data were collected and processed in line with the Data Protection Act 2018.

How are equality data collected?

6.3   The methods used to collect the equality data examined in this research included online, telephone, face-to-face, and paper-based methods – sometimes used in combination; there was also variation in whether the information was provided directly to the organisation by the customer, or gathered and recorded by a 'third party' (i.e. a partner or external organisation, or individual(s) contracted to provide services of some type).

6.4   Research participants also described the use of equality monitoring forms (EMFs)[38] in relation to data collection. This took three main forms: firstly, EMFs formed the basis of the data collection; secondly, EMFs were used separately, as an additional form to accompany the 'main' data collection (note that in some cases these were returned separately); and thirdly, the EMF was developed separately and then integrated with the 'main' data collection.

6.5   The extent to which the collection of equality data is 'required / mandatory' or 'voluntary' is complex and can depend on the rationale for the data collection. For example:

- Where data collections seek to ascertain information in order to award grants, or benefits, or other financial or practical support, interviewees were more likely to say (at least some of) the responses to equality questions were 'required' or 'mandatory'.

- Where equality information was collected to examine differential experiences or impacts of services for different equality groups, the interviewees were more likely to say responses to equality questions were 'voluntary'.

- In many cases, a 'prefer not to say' response option (or similar) was offered as a response to equality questions (see also paragraph 4.17 above). This has implications for understanding what was meant when the collection of equality data was described as 'required' or 'voluntary'. Arguably, **answering the question** may be 'required' but providing information about one's equality characteristics is not.

6.6   These issues are returned to in Section 9.

How are equality data processed?

6.7   The processing of equality data (which covers data input, quality control, data cleaning, data revisions and storage) is highly dependent on the infrastructure and resources which are available to support the particular data collection. Compliance with the

---

[38] An equality monitoring form is a self-contained set of equality questions that can be used in relation to a number of different processes or services.

DPA 2018 is the basic legal requirement, and the organisations in the sample were very conscious of this. For example:

> **Examples**
>
> The organisation is 'very alive to data protection issues' and robust processes are embedded in our systems. We have a raft of data protection policies, and treat the issue very seriously. The information we hold on people is very sensitive, and we are very aware as data controllers of only collecting and holding information that is required to progress a case and carry out our work.
>
> Making sure that data is held securely, following all the legal processes, it's of utmost importance for all our roles.

6.8   The research heard about very varied processes. On the one hand, a number of data collections were identified where the information was gathered on paper (and in most, but not all cases, subsequently input to an electronic system). In other cases, there were very detailed arrangements for quality control and feedback (involving several iterations between respondents and the bodies collecting and reviewing the information), and collections were integrated into extensive IT systems and platforms, with complex and carefully developed data linkage arrangements in place.

6.9   Specific issues referred to in relation to the processing of equality data included:

- Ensuring that personal details (e.g. name and other identifying information) are stored separately from other information in an individual's record

- The requirement to get consent for data holding and sharing, the different ways this was done, and the challenges of doing this in some situations

- Recognising the importance of having a facility to update equality data on a regular / continuous basis in relation to ongoing administrative data collections

- Restricting staff access to data on a 'need to know' basis

- Implementing quality control procedures (e.g. the automated identification of records for deletion based on date information) to ensure that data was not retained longer than was necessary, and

- Recognition that self-reported data cannot be verified.

6.10   Such issues are returned to in Sections 7 and 8.

7 Barriers and challenges in collecting equality data

7.1   This section first discusses the barriers and challenges in relation to the collection (and use) of equality data (in general) and then goes on to provide details of barriers and challenges in the collection and use of sex and gender specifically.

7.2   It is important to note that, given the diversity of the data collections explored (and the relatively limited numbers of organisations and collections that have been included in the research), the barriers and challenges identified are at a fairly general level. In addition, a commitment to participants that no organisation or data collection would be identifiable in the report means that the scope for traditional 'subgroup analysis' is limited.[39] It is important to note that these barriers and challenges did not by any means apply to all organisations / data collections included in the research. However, these issues were raised repeatedly across the range of interviews.

7.3   The main barriers and challenges identified related to:

- The personal and sensitive nature of equality data

- Practical, operational and / or methodological challenges arising from the context within which data collection is undertaken

- Difficulties arising from the complexity of defining equality groups using agreed and standard definitions and terminology, linked to difficulties of 'keeping abreast' of developments in this (fast-moving) field

- Shortcomings in organisational culture, capacity and / or capability.

7.4   Each of these are discussed in greater detail below. These challenges were interrelated and overlapping to a substantial degree, as will be seen from the text.

The personal and sensitive nature of equality data

7.5   The personal and sensitive nature of equality data was said to present challenges for organisations. Participants returned to this theme again and again, expressing the view that applicants / service users / members of the public can be reluctant to provide information about their equality characteristics. Participants suggested that service users may wonder why they are being asked questions which do not directly relate to the service provided; and that some individuals may be concerned that this information might be used to discriminate against them.

7.6   Participants suggested that members of staff can also feel uncomfortable asking people about their protected characteristics because of its personal and sensitive nature – for similar reasons. They may feel the information is irrelevant to decisions about providing

---

[39] Any subgroup analysis (by stratum, say, or by type of data collection) would be based on such small numbers as to be likely to lead to the identification of an organisation and / or data collection.

a service, or that a service user / member of the public may not wish to answer deeply personal questions about their sexuality or religion.

> **Examples**
>
> Equality monitoring can be difficult for people to understand. People don't understand why the organisation is asking what can be regarded as quite personal questions. Even when information is sent out prior to a survey, the level of response can be low.
>
> We are trying hard to increase the diversity of our visitor population. But we have had lots of feedback that asking visitors questions about, for example, their sexual orientation, is off-putting.

7.7   There was also a view among participants that individuals are less likely to disclose personal and sensitive information if they did not trust (i) the organisation asking for the information, or (ii) the rationale presented to them. In some cases (particularly where the collection of the item(s) was mandatory, and when the data collection was an ongoing, administrative collection), discussion focused on the 'best moment' or the 'right time' to collect the equality data. Organisations identified both benefits and challenges of collecting data at the first available opportunity. The main benefit was that early collection of data helped with service planning and offering the most tailored support. However, the challenge was that individuals might find it difficult to understand the rationale for collecting the data before they were fully engaged with the service or policy. Participants explained that where there was an ongoing relationship with a client / service user, it was often possible to collect the information later.

> **Examples**
>
> Some people may be reluctant to disclose a disability or their care-experienced status at application, but they may disclose at a later point, and this would then be added to their record. On disability, specifically, there have been some issues about getting people to disclose at first, for fear of discrimination. Often people will tell us once they are in, but they might not tell us about it beforehand.
>
> Non-disclosure initially increased in the wake of the introduction of the DPA 2018 as individuals were on 'high alert' in terms of what they were disclosing. However, this trend reversed as the DPA bedded in and people were more confident, and trusted that if they disclosed sensitive information, their privacy would be safeguarded.

7.8   One participant also raised the issue of 'intersectionality' (the complex inter-connectedness of people's identities and how this relates to discrimination and disadvantage). This was seen to be an important issue for the 'customers' of one public sector organisation who, it was said, do not see themselves in terms of answers to simple 'tick-box' questions, and as a result they simply refuse to answer the questions, or they spoil the form.

> **Example**
>
> Our 'customers' say that intersectionality is not adequately addressed in these questions and that equality monitoring forms are 'reductive'. These individuals do not want to be pigeon-holed on the basis of these kinds of categories; they want the freedom to define a changing identity which is not 'pin-downable' in the way required by an Equality Monitoring Form.

7.9 More generally, participants talked in terms of 'tensions' in relation to the collection of these data, the 'balances which needed to be struck' and the importance that the collection of equality data should be 'proportionate'. These were ongoing issues for many organisations and were constantly under review.

> **Example**
>
> We do not want to burden people by asking them for a lot of detailed personal information. So, we restrict what we collect to something which is **proportionate** and seen as **essential**.

Practical, operational and methodological challenges

7.10 The challenges described above in terms of the sensitivity and personal nature of equality data were often compounded by practical or operational challenges related to the context in which the organisation is operating and / or methodological approach used to collect the data.

7.11 For example, the social and emotional context and / or the physical environment within which the data collection takes place is not always conducive to the collection of high-quality (complete and accurate) data. Data collection can take place during an emergency (including a medical emergency or a police incident); outdoors when the weather is inclement (which could be the case for visitor surveys); in busy reception areas or other public areas; in places where there is little privacy (in a home with no access to a private space); when an individual has been bereaved; or where a complaint is being made about something which has been the cause of distress. This was a particular issue for organisations providing 'frontline' services where members of the public may be engaging with organisations in difficult or highly charged situations and obtaining equality information could present significant challenges.

7.12 There may also be very limited time available to complete the data collection – particularly if individuals are also being asked a range of other questions at the same time – for example, in a visitor survey, or during a client / patient assessment interview.

> **Example**
>
> In a 15-minute survey, it is difficult to collect everything you want – especially given the situation where you are 'grabbing' people as they leave the place they have been visiting.

7.13   Participants explained that, in some cases, the data may be collected through observation, rather than through direct questioning, particularly in cases where the individual may be unwell, uncooperative, have communication difficulties, or speak another language. In other cases, information about an individual may be provided by a carer, a family member or other relative, or a friend.

7.14   In some public bodies, equality data is collected by a third party (as explained above in paragraph 6.3, this means it is collected by a partner or external organisation, or by individual(s) contracted to provide services of some type). These third parties do not always work to the same requirements as the organisation itself and / or they may not give the data collection the same priority – particularly if the data collection is not compulsory.

7.15   In addition, a few interviewees mentioned situations where the professional approach adopted could mean that the systematic collection of comprehensive equality data was not always seen as a priority. Indeed, collecting equality information – where it was not regarded as essential to the service being operated or meeting an individual's needs – could be seen as not wholly conducive to the client / professional relationship.

7.16   In terms of the challenges of certain methodological approaches, participants described arrangements in which an individual may apply for a service, a place or a benefit online, but then is asked to complete and submit a separate equality monitoring form by post or email. One organisation noted that applicants in such situations do not often complete or return the form.

7.17   Participants explained that language, communication or literacy difficulties – or differences of interpretation – can also prevent or impede the asking and answering of equality questions. For example:

> **Example**
>
> The information captured across the organisation is pretty standard, and the questions asked should be pretty standard. Staff will follow a protocol in taking information from people, but there are likely to be local variations in this. There will also be sharing of 'best practice'. There will also be individual-level variations, how questions are phrased, what form it is recorded on.

7.18   This study also included at least three organisations which collect equality data about **organisations**, not just individuals. For example, in two cases, the organisation provides funding to external organisations and collects information about the *number* of board members and staff in recipient organisations who have certain equality characteristics. Undertaking data collection which relates to organisations rather than individuals is complicated to operationalise, and this information is sometimes not completed well – or indeed, at all.

7.19   Finally, in at least one case the data were collected on paper, but not transferred into an electronic system. This severely limited the extent to which the data could be used.

Lack of standardisation and the pace of change

7.20   There were many examples where organisations described the challenges relating to (i) the lack of standardised approaches, and (ii) the pace of change / the continuous process of review in relation to the collection of equality data.

7.21   Participants in this research highlighted:

- The lack of a standard set of questions to elicit data about equality characteristics

- The lack of a standard set of response categories

- The lack of standardised terminology

- The lack of a standard set of definitions to underpin the collection of equality characteristics.[40]

7.22   While this presented difficulties for individual organisations in their work to gather equality data, it was also pointed out that this lack of standardisation leads to inconsistencies in approach across the public sector, and a lack of alignment between partner organisations.

7.23   This lack of standardisation related to all protected characteristics and socio-economic disadvantage indicators. However, the challenge that was raised most often related to sex and gender. (See paragraphs 7.31–7.32 below.)

7.24   Participants explained that the knowledge of 'what questions to ask' and 'how to ask them' was not routinely in people's skill sets. Therefore, some type of formal guidance / help with identifying standard questions could be useful. However, at the same time, flexibility would be needed to take account of the context in which the organisation is operating.

> **Example**
>
> Different organisations use different formulations, and it would be great to have somewhere you could go for definitive guidance. But in a way, it is not like that, because each organisation has its own reasons for collecting data in particular ways. Disability is a prime example. So many different approaches / definitions.

7.25   Linked to this lack of standardisation was a widely shared view that the collection of equality data was a 'moving target'. Research participants were aware that this was a continuously evolving field (and some were aware of initiatives such as the Scottish Government Chief Statistician's Working Group on Sex and Gender (see Footnote 41 below), and said it could be difficult to keep abreast of the changes which were happening.

7.26   Some examples of organisational perspectives on this are highlighted below.

---

[40] It should be noted that the Scottish Government and NHS Scotland have previously published guidance notes on the use of equality questions. However, these were seldom mentioned by research participants.

> **Examples**
>
> A review of the [] Act is being undertaken. It is thought that, following this, data collection forms will have to be redesigned / updated to align with any changes.
>
> Our survey hasn't particularly kept up with our organisation's transgender policy.
>
> We don't think we are keeping up with gender recognition. Goalposts are moving and it is a political minefield.

Organisational culture, capacity and capability

7.27   There was a range of barriers which participants described in relation to organisational culture, capacity and capability. In relation to organisational culture, examples were described in which:

- Senior management was not focused on, or giving priority to, equality data collection. As described by one contributor 'The equalities remit is often just a small part of somebody's role, and it gets squeezed by other priorities'. Another described the 'organisational inertia' in relation to equality data collection.

- Restructuring and reorganisation were said by some to have absorbed a lot of resources in recent years, with implications for addressing equality data collection issues.

- Senior management was seen as unwilling to invest in equality data collection. The consequence of this was that there was no evidence of the differential impact the service was having on its customers, and consequently no rationale for engaging in equality data collection. It was not clear how this cycle could be broken.

- One organisation was said to focus on issues related to legal compliance only – and did not accept or understand the importance of 'mainstreaming' equality issues.

- An organisation implemented a monitoring system (to ascertain whether the opportunities offered by the organisation were attracting a diverse range of individuals), but it did not take any action to engage in broader issues – not just whether individuals from various equality groups had participated – but whether longer-term outcomes (like their pay, and promotion) were improved as a result.

7.28   In relation to organisational capacity and capability, participants in some organisations said there was insufficient 'infrastructure' available within the organisation to properly support the collection, processing, storage and use of equality data. The kinds of insufficiencies described included:

- The absence of an effective, joined up approach to the collection of equality data with appropriate input from managers, equality and diversity staff, IT staff, analysts, etc.

- Resources not available to undertake data input of paper-based forms

- Insufficient analytical support to maximise the use of the data which had been collected (this point was mentioned specifically in the context of analysing using an 'intersectionality' approach as well as more generally)

- No access to support or advice in designing data collection instruments or protocols for data collection

- Insufficient internal resources to carry out equality data collection resulting in the requirement to use external consultants – who could be very expensive

- An insufficiently well-developed IT platform to enable data to be processed efficiently

- Insufficient resource available [in a partner organisation] to get the extracts / downloads that the organisation needs in a useful format

- Not enough research and analytical capacity / expertise to identify the differential impacts on equality groups.

7.29   (Ongoing) IT development was described as a major issue in some organisations. Participants said that it can require a great deal of both internal and external resources to make changes to IT systems.

7.30   There was a general view that small organisations have more limited capacity to collect data about the equality profile of their users than larger organisations. In larger organisations there could also be challenges – but these tended to relate more to the possibility for 'silo working' and a lack of good communication between, for example, strategic managers, equality leads, analytical staff, administrators, etc.

Barriers and challenges in relation to sex and gender

7.31   As noted above, the sensitivity of the information to be collected, and the lack of standardised approaches presented challenges in collecting data on equality characteristics. However, there was a recurring view among participants that collecting data on sex and gender specifically was especially challenging.

7.32   **Sex** and its relationship with **gender / transgender reassignment**, in particular, was mentioned by a range of organisations. Participants in this research acknowledged that simply offering people a binary choice (Male / Female) when they are asked about their sex (or gender) was no longer appropriate. However, in many cases, IT systems and monitoring forms only provided for these two options. Participants suggested that there was a great deal of confusion about terminology and that the two (different) concepts of sex and gender were often 'conflated'. Indeed, it was common for information about 'sex' to be collected under the heading of 'gender'.[41] (Note that Annex 6 provides more comprehensive information on this point.)

---

[41] Note that the Chief Statistician wrote to stakeholders on 10 December 2020 seeking views on his draft guidance about how to collect, disaggregate and report data on sex and gender. The deadline for submitting views was 12 February 2021.

8 Enablers and facilitators in collecting equality data

8.1   This section discusses enablers and facilitators in collecting and using high-quality equality data. Some research participants described recent efforts undertaken by their organisation to develop and improve their collection and use of equality data. These efforts were spurred both by policies or programmes within their own organisation and by external drivers or factors. Some participants also shared their thoughts about what **could** help their organisation to develop and / or make better use of their equality data collections.

8.2   These experiences and views are discussed below under the following headings:

- Mainstreaming equality

- The importance of being clear about the purpose of collecting data

- The importance of training (for staff and third-party data handlers)

- Data collection methods that work

- Building data management and analytical capacity

- Improving guidance and developing networks for support.

Mainstreaming equality

8.3   Some participants in this study highlighted the importance of their organisation's efforts to 'mainstream' equality. There was no standard approach or model for this, but there was a wide range of suggestions for how this agenda could be advanced. These involved decisions about organisational structures, training and leadership. Participants also emphasised the importance of taking specific steps to change organisational culture and behaviour.

8.4   In some cases, these efforts were tied to external reporting requirements – for example, having to produce statistical returns for government, or demonstrating progress towards outcomes under award schemes.[42]

8.5   In many – if not most – of the organisations included in the research, equality issues were under active consideration and work was ongoing to encourage diversity in decision-making and to reduce inequalities in service delivery. In particular, the requirement for all public sector bodies to produce an Equality Mainstreaming Report on a regular basis reinforces a strong emphasis on the development of equality outcomes.

8.6   Key staff (individually or collectively) often played important roles in championing the equality agenda within organisations. For example, equality and diversity teams, internal equality networks, and research and analytical staff were all said to have been instrumental in previous or ongoing work to improve the collection and use equality data. (Although there was also at least one organisation that had decided **not** to appoint an Equality Officer

---

[42] For example, the Athena Swan programme, or Volunteer Award programmes.

because they wanted to create a culture in the organisation where **everyone** brought an 'equality lens' to the work they are doing.)

<div>

Examples

We have spent a lot of time relaunching our 'Equality Forum' to make sure it really is diverse and that it has real influence in organisational decision-making.

Our Board takes a close interest in looking at the profile of applicants and whether they are representative of the wider population.

</div>

The importance of being clear about the purpose of collecting data

8.7   Section 7 highlighted the perception that some participants had that service users / clients / patients / members of the public did not always understand **why** they were being asked to provide information about their equality characteristics. Some participants suggested that this lack of understanding was often linked to distrust – and could result in unwillingness to provide information about equality characteristics when requested. There was a suggestion that people are **more** willing to respond to personal questions if they see them as relevant, or they see the potential benefit of doing so (if they might get assistance, or if it might have a bearing on their case).

8.8   In the main, organisations that were collecting high-quality equality data with good response rates had invested time and effort in developing clear communication with their customers. For example, they provided a short explanation, in plain English, about why they were asking for this information, what the information would be used for, and how the respondent's privacy in relation to the data would be protected – in line with the requirements and principles of the DPA 2018.

8.9   Organisations that were managing high-quality data collections understood the importance of developing trust with their customers. Part of this process involved explaining to their customers how their data were being used to improve aspects of the organisation's service.

<div>

Example

We carry out a survey among leaders in our area of work. This group used to routinely respond to the equality questions in our survey by writing in 'This is not about gender / disability / race / sexual orientation, etc.' rather than simply ticking the 'prefer not to say' box. However, attitudes seem to be shifting and they understand now why it's important and will disclose their equality characteristics. This is, I think, because more trust has been built around the collection of equality data.

</div>

The importance of training for staff and third-party data handlers

8.10   Organisations that had high-quality equality data collections had also invested time in training staff and third-party data handlers to collect the data. This includes the development of guidance, training materials, and 'scripts' that staff collecting the data can

refer to when speaking to customers. Examples were also given where organisations had introduced training about 'unconscious bias' into their staff induction programmes and staff were given training (and refresher training) specifically about how to collect equality data from their customers.

> **Example**
>
> Work is ongoing to improve understanding of what information should be collected, what it is used for, why it is important, etc. Although the emphasis of this is on improving operational practice and the quality of the interactions with individuals involved, it should also improve data quality. Probationers are being given updated training, and existing staff will be given refresher training.

8.11   Where third-party service providers were collecting the data, payments for services were in some cases linked to completion of the equality data monitoring forms.

> **Example**
>
> Completion of the Equality Monitoring Form is mandatory. It must be completed by all participants. Our external service providers cannot make a financial claim on a programme without the paper Equality Monitoring Form being completed, and the information entered into the system (electronically) for every individual. Individuals have the option to select 'prefer not to say' if they do not wish to disclose their equality information. However, we communicate clearly to participants and service providers that (i) it is important for us to understand who is using our services so that services can be tailored to the needs of the participants, and (ii) it is ultimately beneficial to our customers if we can identify any groups who are under-represented or face additional barriers to accessing our programmes – so that steps can be taken to improve access.

8.12   More generally, there was a recognition that some – perhaps many – of those involved in frontline service delivery were not well versed in relation to social and cultural issues which might be important in the delivery of a service.

Data collection methods that work

8.13   Organisations that were able to collect equality data directly from their customers through an online data collection form generally reported higher-quality, more complete data – although as the example given above indicates, paper completion of forms can also work well where there is a commitment (or requirement) by those distributing the forms to (i) collect them, and (ii) enter them into a database.

8.14   One organisation in the study had recently been through a process to make it mandatory for their customers to complete an equality monitoring form as part of an online application process (which most of their customers were already using). Previously, a voluntary approach had resulted in fewer than 5% completing the form. The new process involved integration of the equality monitoring form into the online application process so that people had to reply to the questions before submitting their application. At the same time, a 'prefer not to say' option was included for all the equality questions. Since this new

process was introduced, the organisation has had a 100% response rate to their equality questionnaire from all online applicants (with only a relatively small proportion selecting 'prefer not to say').

8.15   Some participants highlighted other benefits of collecting data online, including that:

- People with online accounts can update their own information (recognising that people's equality characteristics may change over time)

- Data submitted online is preferable to that collected via an intermediary, or through observation, because of the increased privacy it allows the customer

- It is more cost-effective and efficient to collect data online – there is no time lag and no costs relating to data entry

- Online data collection can be combined with complex data entry checks, which ensure the final dataset is more (logically) consistent.

Building capacity for data management, analysis and reporting

8.16   Section 7 discussed the barriers that some organisations reported in relation to better collection, and making better use of the equality data they collect. However, in many of the organisations, there were ongoing efforts to continuously improve data collection, processing and analysis. The development of new IT systems and platforms incorporating better online data collection facilities, improved data cleaning, and enhanced analytical capability were all mentioned in a wide variety of contexts.

8.17   Some of the public sector organisations involved in this study had large teams of highly skilled data managers and analysts at their disposal. These teams were responsible for producing regular internal and external reports (including reports designated as national or official statistics). Inevitably, organisations with this type of capacity had good systems in place to collect, process, analyse and report their equality data collections.

8.18   There was a range of concrete examples where services or policies had been developed and changed as a consequence of the equality data which had been collected and analysed. This was described as creating a 'virtuous circle' whereby the demonstration of potential benefits to the organisation of collecting and using the data could lead to improved collection and use.

Examples

Previous analysis has suggested that there may be an issue of take-up amongst younger age groups, and so [organisation] has put effort into targeting publicity and literature at younger people.

We are aware of the need to improve how we use the data, how to make the best of it. We are trying to get better at analysing data. With regard to [service area], we need to be able to drill down and look at the success of different interventions / to know what makes a difference.

Improving guidance and developing networks for support

8.19 Organisations with high-quality data collections – and those who wanted to improve their data collections – generally supported the idea of improving guidance and developing networks for support. There was a range of suggestions about how this could be taken forward, including through:

- Better standardisation and coordination of equality data collection – both within and across organisations. This was seen as a key issue in terms of improving the understanding of the experience of different equality groups across the full range of public services. There was frequent reference to the Census 2011 questions, and whether these should provide the standard that all public sector bodies use for collecting equality data. Some organisations said that they were, in fact, using the census questions in their own equality monitoring form. (However, upon examination of the data collections, it was clear that some of these were using variations on the census questions.)

- Establishing and / or participating in equality fora. There is a variety of 'equality fora' in place, both internal and external to organisations, which were reported to be a source of information, support and guidance. These can be important drivers for change. However, not all the organisations participating in this study had the capacity to take part in external equality fora.

- Developing central support and resources. There was a suggestion that it could be helpful to have one (nationally coordinated) 'go to' place for guidance and support.

- Making use of external expertise. Some organisations reported working with other equality groups such as Stonewall and the Equality and Human Rights Commission to improve their approach to collecting equality data.

9 Views on making equality data collection mandatory

9.1   There was a range of views in relation to the question of whether equality data collection should be made mandatory. Whilst some participants were open to this suggestion and / or welcomed a move in this direction, others were more sceptical and had reservations and concerns of various types. Whether participants were in favour or not, a requirement for additional resources was identified, if this option were to be pursued. The main perspectives are elaborated below.

9.2   Participants referred to the DPA 2018 in two main ways in relation to this question as follows:

- There was widespread support for the requirement – if mandatory data collection were to be pursued – of offering a 'prefer not to say' option in relation to individual equality questions. This was seen as vital in order to safeguard an individual's privacy.

- Participants thought that introducing mandatory equality data collection would reinforce the requirement for organisations (as stipulated by the DPA 2018) to be clear about why they are collecting these data, what they will be used for, and why it is therefore in the interest of service users and clients to provide it.

Positive or neutral views towards making equality data mandatory

9.3   In relation to some highly developed administrative data systems (including those that are published as, or contribute to, National Statistics and some other official statistics), there is an ongoing process in place for determining – and reviewing on a regular basis – which equality characteristics should be collected. This process is consultative, and views on the advantages and disadvantages of collecting a specific characteristic are discussed widely among stakeholders. When a 'new' equality item is introduced, it can take time to obtain high-quality data but there are examples of success in relation to this, particularly within the education sector (and specifically related to student enrolment in post-school education). In general, the quality and completeness of the data for mandatory fields is described as very good.

9.4   In some specific cases, making the collection of these data mandatory is regarded as fairly unproblematic as the data are (already) required for operational reasons. In particular, eligibility for apprenticeship schemes, or specific financial support in relation to education or training or benefits would be determined by information relating to age, financial status, postcode (i.e. the individual must live in Scotland), etc.

9.5   A range of participants simply said that if the Scottish Government were to implement this requirement then they (i.e. the organisation) would be obliged to meet it, and they were open to this. They also affirmed the importance of having good equality information about their service users / clients available. This point was sometimes expressed as the Scottish Government providing 'cover' or 'authorisation' to the organisation to implement a change in

their data collection procedures. It was also suggested that a 'phased approach' to any change would help make this manageable.

Concerns and reservations about making equality data collection mandatory

9.6   Some participants expressed concern about making equality data collection mandatory and thought that an 'inflexible' or 'blanket' approach was not helpful. This group was of the view that these kinds of decisions should be made locally by those who understood in detail the requirements to collect (or not collect) specific equality characteristics and who also understood in detail the features of the environment within which data collection takes place. Organisations did not want to be told to collect equality data when they had no evidence that the specific characteristic under consideration was relevant to their service or policy.

9.7   Participants thought that such a requirement would (or could) be disproportionate or inappropriate, or they could not see why it was needed or helpful. They did not see this development as benefitting their own organisation; they saw it only in terms of meeting some (new) requirement from the Scottish Government. Participants suggested that, if the service offered is limited, then the amount of data which should be collected should also be limited; if the service offered is comprehensive then there is more justification to collect a wider range of data.

9.8   In some organisations, the decisions which were currently being made about whether or not to collect specific equality items were grounded in well-developed and robust decision-making processes. For example, if the Equality Impact Assessment in relation to an organisational strategy revealed no impacts on specific protected characteristics, then there was no reason to collect them.

9.9   Additionally, it was argued that any suggestion that providing equality information was 'mandatory' would present particular difficulties and / or would be inappropriate for those services that deal with the public in more challenging situations.

9.10   Some more practical reservations were also raised including that:

- In some cases where – in theory – data collection was already mandatory (i.e. there was a 'mandatory' field in a database), there were 'work arounds' which allowed staff doing the data input to move on to the next field in a database without actually entering a real response (e.g. by putting a letter or symbol in a field).

- It is difficult – perhaps impossible – to enforce completion of equality questions for applications or forms submitted by post.

- Asking for further detail may reduce completion rates for items that are already collected (or may even be a barrier to uptake of services), especially if it was not clear to the client or service user why the information was required.

- Some existing processes and data management systems would have to change, and the implications of this varied across organisations. Some interviewees referred to in-

house systems which could be readily adapted as required, while others had bought-in systems which nevertheless offered good scope for local configuration by in-house staff. In contrast, other interviewees spoke of complex IT systems involving external suppliers and ongoing contractual and oversight arrangements that meant that changes could be more complicated to achieve and incur financial cost.

- Some existing equality data collections have been developed through extensive consultation with service users. A requirement for mandatory data collection could undermine this activity.

- Since the information is 'self-disclosed' it cannot be validated. Making it mandatory will therefore not necessarily improve the quality of the data.

The requirement for additional resources and support

9.11   Irrespective of their views on whether or not making (more) equality data collection mandatory was desirable, participants identified requirements for additional resources and support to be made available if this option were to be pursued as follows:

- This would need to be accompanied by an initiative (perhaps a Scottish Government led public information campaign) to explain the reasons why it had been adopted. As has been alluded to earlier, members of the general public or service users or grant applicants do not necessarily understand or agree with the requirement to disclose sensitive personal data.

- There would need to be additional support including IT support (for amending and upgrading IT systems), and also enhanced capacity in relation to IT governance, analysis, and reporting.

9.12   Finally, participants sometimes linked a move towards making (more) equality data collection mandatory with a move towards more standardisation of data (which in turn would require more resources and support).

10 Reflections on 'what works best'

10.1   This final section sets out the reflections and insights in relation to 'what works best' in equality data collection. The section draws on material presented in earlier sections of this report.

10.2   The research identified four key themes in relation to 'what works best' as follows:

- The **mainstreaming equality** agenda has provided substantial impetus to the issue of equality data collection.

- Clarity about **why** questions are asked and **how** the information will be used is essential to establishing and improving high-quality equality data collection.

- The extraordinary **diversity of the data collections** under consideration means that it is vital to take a tailored (case-by-case) approach to improving the collection of equality data.

- There is a need for a **'go-to' place** to help with the 'nuts and bolts' of equality data collection, especially for organisations where less in-house expertise is available.

10.3   Each of these themes is briefly elaborated below.

Mainstreaming equality

10.4   The mainstreaming equality agenda – including the requirement for regular progress reports – has been an important driver in promoting a focus on, and improving, equality data collection. There were varied accounts of how, specifically, this agenda was translated in relation to individual organisations.

10.5   A range of organisations had responded to the agenda of mainstreaming equality by developing a positive organisational culture and proactive leadership, investing in infrastructure, and ensuring that there was a joined up, cohesive approach to equality data collection.

10.6   In addition, as part of this agenda, there was an appetite for more guidance and greater consistency in the collection of equality data. It was thought that the Scottish Government had a vital role to play in leading on this.

Clarity about why questions are asked and how information will be used

10.7   There was compelling evidence from a range of participants about the importance of ensuring that all those involved in developing equality data systems were clear about why questions are asked and how the information which is gathered will be used. As set out at various points in the report, this clarity is already a requirement of the DPA 2018. However, if equality data collection is to be improved, the need for clarity goes far beyond a (narrow) legal requirement.

10.8   Organisations thought that this clarity helped (or would help) to build trust with those from whom the information is being sought, and gives (or would give) them confidence that their data will be used to benefit both themselves and others.

10.9   This clarity is often prompted by – and developed as a consequence of – 'normalising' and 'mainstreaming' the equality agenda within organisations (see above). Thus, finding ways to encourage this is one of the keys to improving equality data collection.

The diversity of data collections and the requirement for a tailored approach

10.10 The range of data collections encountered in the research was vast. The collections were highly diverse in terms of subject matter – as expected. But they were also diverse in terms of their operational contexts. This variation was seen in terms of (i) the physical locations in which data collection took place, (ii) the emotional and social contexts in which data were collected, (iii) the individual(s) carrying out the data collection, and (iv) (partly as a consequence of (i) to (iii) above) the methodological approaches to collecting the data.

10.11 This means that any approach to equality data collection must be very carefully tailored to the situation at hand. The task of improving equality data collection cannot take a 'one size fits all' approach.

10.12 The research found (and this is perhaps the one general finding) that there are many advantages to collecting data online. These advantages – whilst not universally achieved – tend to mean that an online collection can support enhanced rates of disclosure and improved data quality because (i) the absence of an intermediary offers better privacy to individuals providing the information, (ii) individual records can easily and continuously be kept accurate and up-to-date, and (iii) the opportunity for (sometimes highly complex) data entry checks to be implemented. However, online data collection is not possible in all cases.

A 'go-to' place for organisations with less specific equality data collection expertise

10.13 Equality data collection is a highly complex undertaking. A range of larger organisations (some of which were involved in equality data collections that had been designated as 'National Statistics') talked about the internal and external networks and sources of advice they could access on an ongoing basis. By contrast, a range of other organisations, including, but not exclusively, smaller organisations, said that they felt there was a lack of a 'go-to' place for advice and expertise on equality data collection, within their own organisation or externally.[43] This was notwithstanding that many of these organisations belonged to some kind of 'Equality Network' which they found helpful.

10.14 Many of the questions participants had were about the 'nuts and bolts' of equality data collection. For example, how should the question be asked? What response

---

[43] See Footnote 40 above.

categories should be offered? How should you define ethnicity or religion? What are the best ways of presenting and interpreting data involving small numbers?

10.15 In some instances, organisations said that they found it hard to keep up with the developments in equality data collection, and they were conscious of the ever-shifting requirements.

Annex 1: Participating organisations

- **Health and Social Care Organisations**
  - Care Inspectorate
  - NHS Fife
  - Mental Welfare Commission
  - Scottish Ambulance Service
- **Employment, Labour and Welfare Organisations**
  - Highlands and Islands Enterprise
  - Independent Living Fund
  - Social Security Scotland
- **Education and Skills Organisations**
  - Forth Valley College
  - The Open University in Scotland
  - Scottish Funding Council
  - Skills Development Scotland
  - Student Awards Agency Scotland
  - University of St Andrews
- **Justice Organisations**
  - Police Investigations and Review Commissioner
  - Police Scotland
  - Scottish Children's Reporter Administration
  - Scottish Criminal Cases Review Commission
  - Scottish Prison Service
- **Local Authorities**
  - Argyll & Bute Council
  - Edinburgh City Council
  - Highland Council
  - North Lanarkshire Council
  - Midlothian Council
- **Other**
  - Cairngorms National Park Authority
  - Creative Scotland
  - Historic Environment Scotland
  - **sport**scotland

Annex 2: Data collections included in the research

Stratum	Name of data collection
Education and skills	Student Record Template
Education and skills	Student Record Template – Higher Education
Education and skills	Student Enrolment Template – Further Education
Education and skills	Application for student support
Education and skills	Application for Disabled Students Allowance
Education and skills	Student Record Template
Education and skills	Funding application form
Education and skills	Hate Incident Monitoring Form
Education and skills	Apprenticeship applications
Education and skills	Student Record Template
Education and skills	Equality, diversity and inclusion survey
Education and skills	Careers Information Advice and Guidance Customer Feedback
Employment, labour and welfare	Feedback and Outcome Monitoring - Programme and Event Participants
Employment, labour and welfare	Transition Fund application – Equalities Monitoring Form
Employment, labour and welfare	Benefits applications
Employment, labour and welfare	Client satisfaction survey
Health and social care	Experience of Living in a Care Home
Health and social care	Local Visits
Health and social care	Emergency Detention
Health and social care	Needs Assessment for Patient Transpport
Health and social care	A&E Admissions
Health and social care	Community First Responders
Health and social care	A&E Attendances
Justice	Application for complaint review
Justice	Application for case review
Justice	Interim Vulnerable Persons Database
Justice	Crime and Incident Management System
Justice	Prisoner Survey
Justice	Visitor Survey
Justice	Prisoner record system

Justice	Case management and referral system
Local authority	Housing Options
Local authority	Homeless Applications
Local authority	Consultation on 'How Good Is Your Place?'
Local authority	Standard Membership Registration
Local authority	NHS Referral Registration
Local authority	Customer Satisfaction Surveys
Local authority	Community Justice client data
Local authority	Pupil record template
Local authority	Child Assessment Information
Local authority	Adult social care client database
Local authority	Annual Student Survey
Local authority	Enrolment data
Other public sector body	Applications for grants (various types)
Other public sector body	Survey of Scotland's Screen Crew
Other public sector body	COVID-19 Impact and Recovery Survey
Other public sector body	Volunteer Programme
Other public sector body	Visitor Survey
Other public sector body	Volunteer Equalities Monitoring Survey
Other public sector body	Wellbeing Survey
Other public sector body	Consultation on 'What's Your Heritage?'
Other public sector body	Visitor Surveys
Other public sector body	Active Scotland Outcomes Framework – Schools survey
Other public sector body	Active Scotland Outcomes Framework – Club survey
Other public sector body	Investment Applications

Annex 3: 'Audit' of equality data – template 1

Item / Descriptor / Question	
Organisation (Name)	
Data Collection ID	
Name of Data Collection	
Description of Data Collection	
Type of Data Collection (Admin / Survey / Census)	
Data collection method(s)	
Frequency of collection (e.g. ongoing, regular intervals, intermittent, one-off)	
Geographical coverage	
Population coverage	
Is the data collection published? (y/n) (Please explain)	
Smallest unit of analysis for purposes of reporting (e.g. national, NHS board, local authority, etc.)	
Protected Characteristic – Age (y/n)	
Protected Characteristic - Religion / belief (y/n)	
Protected Characteristic – Race (y/n)	
Protected Characteristic – Disability (y/n)	
Protected Characteristic – Sex (yes/no)	
Protected Characteristic - Sexual orientation (y/n)	
Protected Characteristic - Pregnancy and maternity (y/n)	
Protected Characteristic - Marriage and civil partnership (y/n)	
Protected Characteristic - Gender reassignment (y/n)	
Postcode (y/n)	
Socio-economic disadvantage(y/n)	
Other, please specify:	

Annex 4: 'Audit' of equality data – template 2

Organisation (Name)	
Data Collection ID (1-5)	
Name of Data Collection	

For each protected characteristic (age, religion and belief, race, disability, sex, sexual orientation, pregnancy and maternity, marriage and civil partnership, gender reassignment) or socio-economic disadvantage, **please provide:**

- Question(s) used to elicit information (including any definitional issues), response categories and
- Response rates for individual questions (in last published report, or during a time period stated by the interviewee)

Protected characteristic	Question(s)	Response rate / period
Age		
Religion & belief		
Race		
Disability		
Sex		
Sexual orientation		
Pregnancy and maternity		
Marriage and civil partnership		
Gender reassignment		
Postcode		
Socio-economic disadvantage		
Other – please specify		

Annex 5: Topics for qualitative interviews

Organisation (Name)	
Data Collection ID (1-5)	
Name of Data Collection	

Equality Data ARE Collected

1. How does the organisation use these data?

2. What, if any, challenges have there been for the organisation in collecting these data?

3. Do you have any views on the quality and completeness of the data? What reasons would you attribute to low (or high) response rates in relation to these data?

4. What, if any, actions are being taken by the organisation to monitor, develop and / or improve the collection of these data?

5. Would you have concerns about collecting these data in the future? Would you have any concerns if data collection were made mandatory for these data in the future?

6. To what extent – if any - have issues related to data protection / GDPR affected decisions in the organisation about the collection of equality data?

Equality Data ARE NOT Collected

1. What factors have informed the organisation's decisions not to collect these data for this particular collection?

2. Are there any plans within the organisation to collect these data in the future? What would need to change (or what assistance would be needed) to make this possible?

3. Would you have any concerns if data collection for these data items were made mandatory in the future?

Any Other Comments

Do you have any other comments in relation to the collection, processing, use, or development of equality data within your organisation?

Annex 6: Collection of information on individual protected characteristics

An overview of the questions, definitions, response categories, and variation on these, in relation to each individual protected characteristic is presented below.

Age

- The collection of age information was seen as relatively uncontentious by interviewees, with some saying that people were very used to providing this information in different contexts.

- Information was gathered in four different ways by asking for either (i) age (ii) age group (iii) date of birth or (iv) year of birth.

- Most commonly individuals were asked to provide their date of birth, thus offering an accurate and dynamic indicator of age. The second most common way of gathering age information was by asking people to indicate an age range to which they belonged. There was variation in the age range options offered between and within organisations. However, it was often clear that the options offered were linked to the focus of the relevant service – for example, in the case of age-related eligibility for services.

Religion and belief

- Most often questions asked about 'religion', with no reference to 'belief', but in a few cases, questions asked about 'religion or belief'. One organisation asked two separate questions, one on religion and one on belief.

- Questions were often similar to that used in the 2011 census question – in terms of both question wording and response options – although slight variations were also common (e.g., 'Christian –Protestant' instead of 'Christian – Church of Scotland').

- Questions typically asked 'What religion, religious denomination or body do you belong to?' Other forms of question included: 'What best describes your religion or belief?', 'What is your religion?', and 'Which group do you most identify with?

- Response categories typically included: Buddhist, Christian – Church of Scotland, Christian – Roman Catholic, Christian – Other, Hindu, Jewish, Muslim, Sikh, Prefer not to say. Additional / alternative Christian denominations were offered in some cases. There was also occasional inclusion of 'Pagan', 'Spiritual', and 'Bahai'.

- There were a number of different variations of 'no religion' offered by organisations, including 'none', 'no religion', 'non-religious (atheist, humanist etc)', and 'atheist, agnostic or no religious affiliation'.

- The number of response categories varied – one organisation offered six response categories, while a small number of organisations offered very extensive lists of categories.

Race

- There was variation in the questions used and response options offered both between and within organisations. However, questions often appear to be modelled on the 2011 census question which uses response options structured around six main categories and sub-categories within those (although with wide range of variations also apparent in the detail of the questions).).

- Where information on questions / field names was provided, these asked about 'ethnicity' rather than 'race' – as is the case with the 2011 census question.

- There was variation in terms of phrasing and terminology, with examples including the following: 'What is your ethnic group?' (i.e. the question used in the census), How would you describe your ethnicity?', 'Which of these groups do you consider you belong to?', 'What best describes your ethnic group?', 'Which of the following best describes your ethnic group or background?'

- The following (different) terminologies were all used: ethnicity, ethnic group / sub-group, ethnic background, ethnic origin, national identity.

- There was a mix of approaches in term of using a single 'layer' of response options, or using main categories and sub-categories.

- While there was a lot of commonality in the response options offered, there was also a great deal of (often slight) variation. The number of options ranged widely (e.g. six options used in one case; over 70 used in another case). There was also variation in the use of 'please specify' options, and the wording of 'prefer not to say' options.

- There was some limited use of multi-part questions – e.g., asking about ethnic origin *and* national identity.

Disability

- The information gathered on disability was wide ranging and often reflected the needs and practicalities of different services – i.e., the information was often used to ensure that the needs of individual service users could be properly met. As such the question and response options were also very varied with some organisations simply aiming to determine if an individual had a disability or long-term condition, and others gathering very detailed information on the nature of a person's physical and mental health and the impact on their daily functioning.

- Two-part questions were common, with individuals first asked if they had a disability – with some either directly referencing or alluding to the definition in the Equality Act – and then asked for further details of any disability or long-term condition, either by using a list of tick-box option or inviting a free-text response.

- Those using a single part question included those asking a simple yes / no / prefer not to say question as to whether somebody has a disability, and others offered a list of different disabilities / conditions for individuals to respond to.

- There was also some limited use of separate observed and self-defined ethnicity.

Sex

- In all the cases where a specific question was provided to the research team, the term 'gender' rather than 'sex' was used. For example, questions asked: 'What is your gender?', 'What is your gender identity?', 'What best describes your gender?'

- Interviewees were often aware of this as an issue, with one specifically referring to sex and gender being 'conflated' in their collection of equality data.

- In more than a third of cases, respondents were (as with the 2011 census question) offered a binary response choice of Male / Female, with or without a 'prefer not to say' option. Again, there was widespread awareness that this was not appropriate, or did not reflect current thinking on this issue.

- Where additional response options were offered, these commonly included: 'Other' (with or without the option to specify) and 'Non-binary'. Other less common response options included: 'man / woman', 'intersex', 'gender fluid', 'other gender identity', 'I identify with another term', 'In another way', 'I prefer to self-describe', 'unknown', 'indeterminate'.

- In three case, response options specified the inclusion of trans individuals: 'Male (including trans man) / Female (including trans woman)', or 'Male (including transgender) / Female (including transgender)'; in another case there was an additional question asking if the respondent had ever identified as transgender.

- There was a small number of examples of organisations using 'title' (Ms, Mr etc) as a proxy for gender.

Sexual orientation

- Questions on sexual orientation were formulated in a variety of ways, but most commonly asked people how they would 'describe' their sexual orientation: for example: which of the following best describes your sexual orientation / how you think of yourself / how you think of your sexuality. Questions generally used the term 'sexual orientation', although there was occasional use of the word 'sexuality'.

- The response options commonly included the categories of heterosexual / straight, gay man, gay woman (or a combined category of gay man / gay woman), bisexual, other. There was some variation in the wording used, most notably in whether the term 'lesbian' was included and whether 'straight' was included alongside 'heterosexual'. Occasionally, additional options (e.g., 'asexual', 'in another way', 'not sure') were offered.

- There was one example of an organisation using a different style of descriptor for the response option offered as follows: 'Please describe your sexual orientation: Towards people of a different sex (straight) / Towards women of the same sex (lesbian) / Towards men of the same sex (gay) / Towards people of both sexes (bisexual) / Other (e.g., asexual).'

- Some response lists included an option for individuals to self-define or specify further if they selected 'other'.

Pregnancy and maternity

- Information on pregnancy and maternity was collected in a few cases only.

- The information gathered in the data collections studied mainly focused on 'pregnancy', rather than 'maternity' (defined in the Equality Act as the 26 weeks following birth). Additionally, this information was sometimes collected on a household rather than individual basis ('Is anyone in the household pregnant at the moment?'), reflecting the information needs of the service being provided.

- In a number of cases, interviewees said that while information on pregnancy / maternity was not systematically recorded in a format suitable for quantitative analysis, it would be included in case notes in a narrative form where this was relevant, or (in a few cases) they said that a separate record may be created for an unborn child.

Marriage and civil partnership

- Information on marriage and civil partnership was only collected in a handful of cases.

- Questions varied in terms of whether they asked about an individual's 'legal status' or asked how an individual 'thinks about themselves'. Some questions referred to *marital* status only while others referred to *marital or civil partnership* status. However, all offered civil partnership response options.

- Response options offered ranged from three categories (married / civil partnership / other, plus 'prefer not to say' in some cases) to lengthier lists including single, co-habiting, separated, divorced, widowed and the civil partnership equivalents.

Gender reassignment

- Only in a small number of cases did the data collections identified by the research gather information about gender reassignment.

- Generally speaking, questions asked about 'identity', asking if individuals identified (or had ever identified) as trans / transgender / a transgender person. One question asked if the individual had undergone or was undergoing gender reassignment, and two questions asked if an individual's gender identity was the same as the gender they were assigned at birth, or if their current gender (sex) was the same as they were born with.

- Although questions commonly used the term 'gender', both 'sex' and 'gender' were used in a couple of cases.

- A two-part question was used in one case ('Is your current gender (sex) the same as you were born with?' 'Do you identify as trans?').

Socio-economic disadvantage

- In almost all cases, organisations gathered postcode information. However, this was generally collected for administrative reasons and was not routinely used as an indicator of socio-economic disadvantage.

- Where organisations gathered information on income – and this was only reported in 13 cases – this was always because there was a specific need for this information related to service eligibility.
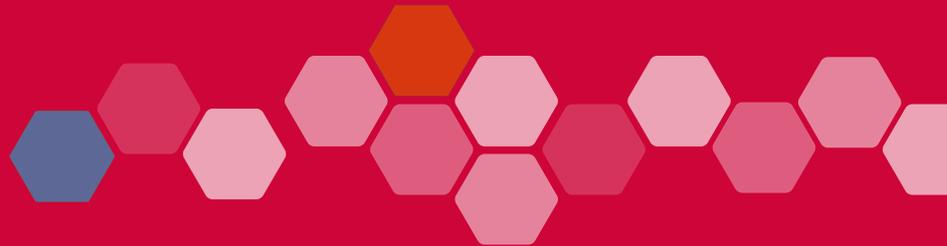
social research