UK Statistics Authority

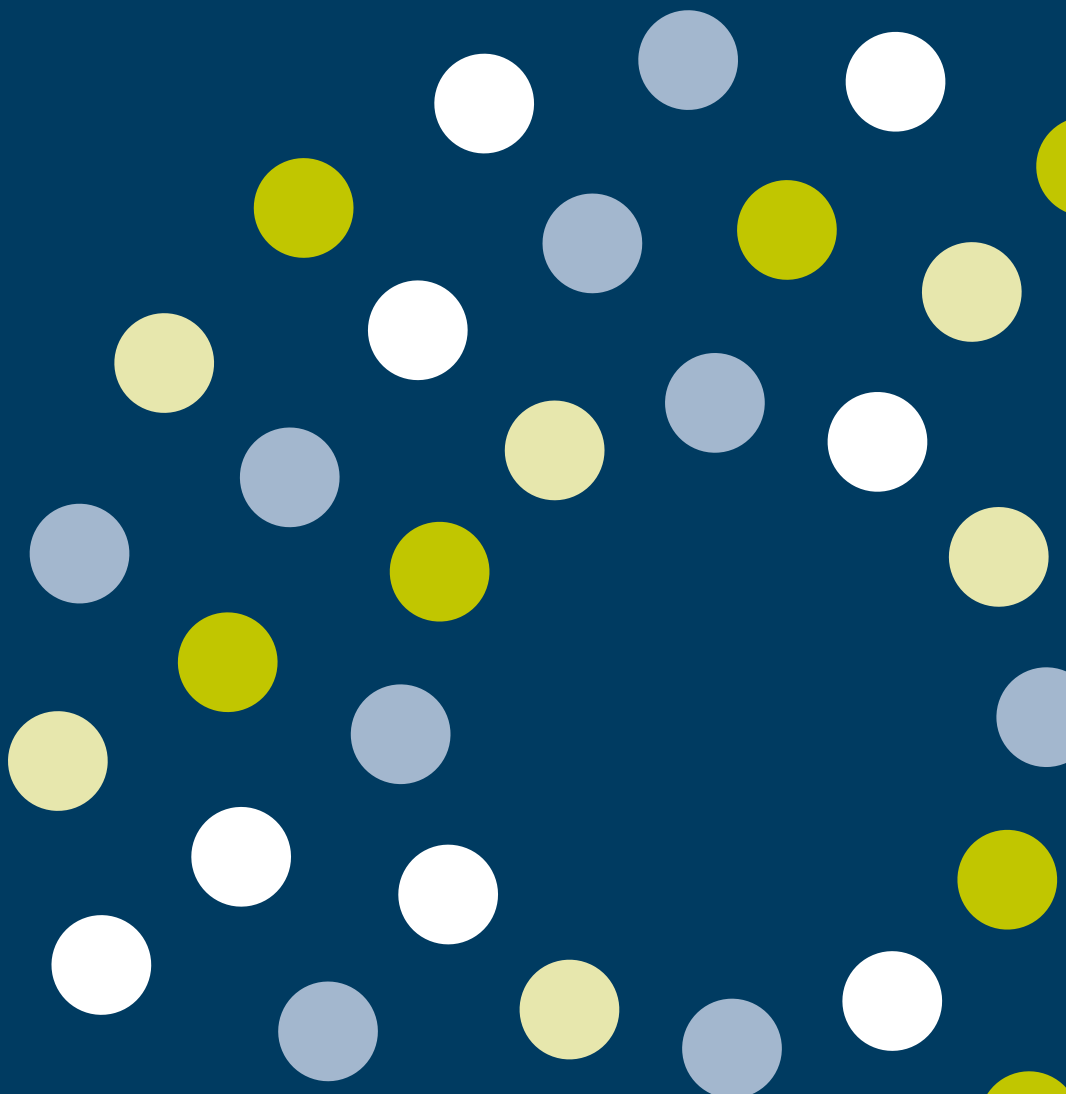
In partnership with ROYAL STATISTICAL SOCIETY
DATA | EVIDENCE | DECISIONS

Independent Report on the 2025 UK Statistics Assembly

Prepared by:
The National Statistician's Expert User Advisory Committee

Chair: Professor David J Hand
d.j.hand@imperial.ac.uk

4 March 2025



Contents

1. Introduction

From 2023 to 2024 Professor Denise Lievesley led a review of the UK Statistics Authority (the Authority) within the framework provided by the Cabinet Office Public Bodies Review Programme. The central conclusion of that review was:

> "It is time for the Board to move into a more visible, ambitious space, primarily through establishing a Triennial Statistical Assembly which will consult widely with statistics users and producers to understand the range of views regarding the priorities and data needs for the UK."

The intention was that:

> "This will lead to a more transparent and robust setting of the statistical agenda with a greater emphasis on user needs."

The first of these Statistical Assemblies (henceforth 'Assembly'), jointly organised by the Authority and the Royal Statistical Society (RSS) and chaired by Professor Cathie Sudlow, took place on 22 January 2025. It sought to bring together the widest possible range of users and producers of official statistics to discuss the priorities, opportunities and challenges facing the statistical system in the short to medium term. That system encompasses the Office for National Statistics (ONS), the wider Government Statistical Service (GSS), other UK-wide and devolved public bodies that produce official statistics, and indeed other sources which feed, or could feed, into official statistics.

The aim of this report is to present the points made during the Assembly, including to:

- highlight user needs and priorities, as well as opportunities and trade-offs, flag important data gaps, and where series are not needed
- recommend how system priorities could align with these needs
- inform delivery planning for the ONS, Office for Statistics Regulation (OSR), GSS, and other statistical producers
- advise on where and how those outside government could best align
- encourage cross-sector discussion
- contribute to the Authority's next five-year strategy

Of course, the extent to which users' wishes and needs can be met is dependent on resources, and, although that is a matter beyond the remit of this report and of the Assembly, it is necessary to be mindful of the constraints on public finances and the need for the statistical system to prioritise and make trade-offs.

The report is structured as follows.

- Section 2 describes the format of the Assembly.
- Section 3 describes priorities which emerged from the meeting. In particular, Subsection 3.1 presents the main priorities at the highest level, while Subsection 3.2 presents suggestions which could lead to improvement on a short time scale.
- Section 4 describes issues which emerged with some consistency across the separate sessions.
- Section 5 presents summaries of the conclusions and recommendations of each of the fifteen breakout sessions. Within each of these a handful of points have been emboldened as being of particular importance.
- Section 6 describes the key points made in a plenary session on the UK statistical system as seen from an international perspective.
- Section 7 describes some recommendations for future assemblies based on what was learned from conducting this one.

2. The Assembly

A call for contributions to the Assembly event was widely publicised, and an Assembly Delivery Group including representatives from a range of sectors designed a schedule intending to cover as many as possible of the concerns. [1] This consisted of four plenary sessions and fifteen breakout sessions on particular topics, grouped into three blocks of five.

The plenary sessions included welcomes from the UK Statistics Authority Chair Sir Robert Chote and the RSS President Sir John Aston, and an introduction from Dame Kate Barker from the Assembly Delivery Group, who also commented on how the session topics were selected. A presentation on international perspectives involved Steve MacFeely (Head of Statistics, OECD), Francesca Kay (Assistant Director General, Central Statistics Office, Ireland), and Vipin Arora (Director of the US Bureau of Economic Analysis). The final plenary presented immediate feedback from the breakout sessions by Professor Sudlow; and a brief description of the next steps by Professor Hand.

The 15 breakout sessions covered the topics: Economic statistics and measuring progress; Coherence; Data sourcing, quality and methods; User engagement; Health disparities; Public sector performance; Data linkage; Net zero; The future of the Census; Business statistics; Equalities; Artificial intelligence (AI) and technology; Crime; Labour market; and Local and regional data.

863 people registered to take part, 556 attended on the day, with 330 in-person and the remainder online. A breakdown by sector is presented in Table 1.

Table 1: Breakdown of attendees by sector.

Sector	Number
Academia	47
Advisory panels	7
Arm's-length bodies	191
Business and industry	52
Charity	26

[1] Organisations represented on the Delivery Group are: Academy of Social Sciences, British Chambers of Commerce, Government Statistical Service, Joseph Rowntree Foundation, National Statistician's Expert User Advisory Committee, Northern Ireland Statistics and Research Agency, Office for Statistics Regulation, Royal Statistical Society, Scottish Government, UK Research and Innovation, UK Statistics Authority, Welsh Government, Westmoreland and Furness Council.

Table 1 continued: Breakdown of attendees by sector.

Sector	Number
Devolved governments	32
Government departments	102
International	10
Local authorities	37
Media	4
Members of the public (or other)	15
Public services (such as housing, NHS trusts)	7
Think tanks	26

98 Assembly attendees completed a post-event evaluation survey. From that, the key findings include:

- the average Assembly satisfaction rating (on a scale of 1 to 5) was 4.2
- 99% felt the Assembly met its aims
- 92% valued the plenary sessions
- 79% rated the breakout groups as good or excellent
- average satisfaction with the diversity of discussion themes (on a scale of 1 to 5) was 4.3
- 95% were likely to attend future Assemblies
- 69% said the Assembly enhanced user engagement

This report from the Assembly was prepared to a tight timetable, with the aim of feeding into the Authority's preparation of its next five-year strategy. An initial draft was prepared by a subgroup of the National Statistician's Expert User Advisory Committee (NSEUAC), which was refined by the entire NSEUAC. Input and clarification was also obtained from Professor Lievesley and Jonathan Everett (RSS).

3. Priorities

3.1 High level priorities

Underpinning the design and format of the Assembly was the wish that discussion should embrace both needs and value. Needs, that is specific user requirements, are of course important and should be heard. Value comes from addressing needs, and impacting on the public good, through advances in a modern statistical system. As already noted, all of this is dependent on resources.

The following four high-level priorities are therefore suggested for the Authority and the GSS to pursue and to help frame their decisions and actions in response to the Assembly.

1. Re-invigorate sustained and effective user engagement, in which official statistics producers take a lead in understanding the needs for statistics and curating relevant sources, to help answer the questions that the public, businesses, local government, the media and academics, as well as policy-makers have about the economic, social and environmental situation. This would also help understand and increase the value of statistics.
2. Develop a portfolio of official and unofficial sources, along with use of appropriate methodologies, to ensure user needs for more granular statistics are met (small areas, urban/rural, sub-groups of society, under-represented groups, and so on).
3. Commit to, invest in, and take a leadership position in a significant scaling up in the use of administrative data, as well as improvement of its quality and coherence, across the entire portfolio of sources of official statistics, including government departments and external bodies, alongside and integrated with surveys, censuses, and other types of data.
4. Recognise the needs for UK-wide statistics and advocate for, and support, harmonised data where desirable.

Since the role of NSEUAC is to "provide strategic advice and insights to the National Statistician about how the UK statistical system can best maximise effective engagement with all users to enable the UK Statistics Authority and the wider system", we urge the Authority to continue to raise public awareness of the existence of this committee and how to contact it.

3.2 Immediate actions

The following are suggestions for actions that could be taken quickly and, we believe, without significant cost, or which could lay the foundations for later decisions and so expedite them. The numbers at the end of each suggestion relate to the numbering system in Section 5, and follow the order of sessions in the Assembly programme.

1. Establish a forum of users and producers to develop a plan for better coherence across the statistical system. In particular, the forum should decide what could be achieved quickly. (5.2.2.)
2. Engage official statisticians more in the design of digital data architectures for administrative data from diverse sources. A cross-GSS group could make a start on this, building on and coordinating existing departmental efforts, to enhance standardisation, in order to feed into more sharing. (5.3.2, 4, and 9.)
3. The challenge of recruiting statisticians. Ask the RSS, ONS, Market Research Society, National Centre for Social Research (NatCen), Ipsos, Kantar, and others, to form a working group to consider and recommend how methodologists and statisticians could be better developed and retained. Perhaps an Academy or specific career development pathways might be appropriate mechanisms. Establish a graduate recruitment programme along the lines of Ministry of Defence sponsorship of students into relevant degrees. (5.3.17.)
4. Provide guidance on statistical methods for coping with discontinuities in time series arising from changing data collection methods. An academic review could be set up quickly. (5.3.16, 5.9.10.)
5. All online published government statistics should include a ‘Comments’ box, permitting user feedback and better understanding of who the users are. (5.4.2.)
6. Publish good stories on how data are used for public benefit. (5.4.19.)
7. Enhance communication efforts to clearly articulate the direct benefits of data linkage and sharing to the public. (5.7.1.)
8. The ONS or GSS should establish an online ‘trust centre’ on the website to provide transparency about data usage and security protocols, similar to the models used by Stats Canada and CSO Ireland. (5.7.16.)
9. Develop the GSS Environment, Climate and Nature theme to include user forums and events, bringing together producers and users of statistics relating to the theme, and helping build a diverse user community for the theme and a tailored calendar of engagement activity. (5.8.7.)

10. Publish detailed information on the use of administrative data in the last census, including areas where response rates were lower and how administrative data was used to fill gaps. (5.9.17.)
11. Set up a forum (including business groups) to conduct research and recommend a communication strategy to explain the importance of data collection to business owners and encourage their participation. (5.10.13.)
12. Transition from reporting average weekly earnings to average monthly earnings to better reflect current business practices. (5.10.15.)
13. Produce coherent definitions of ethnicity, family types, and areas with inconsistency, and require their use across all government departments. (5.11.1 and 2.)
14. Produce a standard document explaining why ethnicity information is required and why it is important. (5.11.9.)
15. Produce a dashboard showing how different crime rates are evolving, and showing an impact-weighted overall measure. (5.13.1 and 12.)
16. Promote collaboration among local government organisations to share resources and develop common data models. Develop consistency across local authority data systems, especially for administrative data. (5.15.7.)

4. Common issues

The 15 breakout sessions covered a diverse range of different topics, but some clear common themes emerged. These include the following.

1. The need to enhance agility and flexibility in response to a rapidly changing world.
2. Take advantage of technological advances, such as AI. AI could be applied in various roles, such as improving analytical pipelines, automation (of data capture and elsewhere), coding and coding checking, and so on.
3. Make more use of administrative data, explore how to continue to gain insights from surveys, and in general actively look for alternative sources of data in order to meet user needs.
4. Focus attention on the quality and consistency of administrative data.
5. The merits of automating data capture as much as possible, via automatic surveys and direct access to business and other administrative data. This will require the involvement of statisticians in the design of data architectures and training for frontline people involved with data capture.
6. Accelerate and improve the capacity for the linkage of data, and improve coherence and compatibility. Ensure consistent data standards and definitions across departments, regions, and nations for effective comparison and integration.
7. Improve the understanding of what data are needed at a local level, and improve its quality in granularity and timeliness.
8. Work to improve data capture for under-represented groups and build a consensus on how to represent ethnicity in data.
9. Model changing work-patterns and develop more holistic measures of work, including volunteering, carers, and others.
10. Ensure a respectful, constructive and engaged feedback loop between producers and users to address transparency and need.
11. Improve access to micro data for research purposes.
12. Improve cross-governmental data sharing.
13. Better integration across the GSS. For example, some departmental statistics could benefit from ONS input and vice versa. This would also materially aid standardisation.
14. Develop effective performance metrics so that policies can be monitored, including for public sector and user engagement.

15. Enhance transparency and take a more active role involving users and producers to address public concerns about data sharing, trustworthiness, and ethical use and to make the case for the value of data. In part, demonstrating trustworthiness means producing statistics that are more relevant to individuals and developing measures that connect to public perceptions (for example, on crime and economics).
16. Improve user and public engagement and communication. Among other things this means broadening the groups of users that are routinely engaged, effectively communicating research findings to decision-makers and the public to inform decisions and demonstrating to the public the value driven by their data.

5. Key points and recommendations of each session

Each session at the Assembly involved an interactive, wide-ranging, detailed and open discussion, specifically designed to elicit user and producer opinion in depth, and reflecting the breadth and depth of experience, the diversity, and the mix of interests of the attending delegates. A very large number of contributions were made, although some sessions could have benefitted from being longer.

This section identifies the main issues raised within each session: some of which are important observations on the statistical system as it is, which need to be discussed; some are opinions of long term future and idealistic requirements; and others are practical recommendations. They are grouped into themes appropriate to each session, and are left as far as possible as they were expressed in discussion to reflect the specific nuance of the individual comment or issue. Some points are fairly subtle, and occasionally technical, but could have a sizable impact on and be beneficial for the Authority. These might not be the 'headline' issues, or the most popular issues, but might have significant impact if developed.

In some cases, the answers to the questions posed by delegates already exist within the GSS (such as, “should collect data on X”, “should tell public about X”, and so on) but the fact that such an issue has been raised indicates a lack of communication somewhere in the system.

While it is the remit of the Authority to decide which of the following issues to tackle, we have sought to identify a handful within each subsection which we regarded as most pressing, impactful, or needed. These are indicated in **bold**.

5.1 Economic statistics and measuring progress

> What needs to be achieved to ensure the measurement of economic statistics keeps pace with societal and technological changes?

Methodological

5.1.1 Enhance agility and flexibility. Agility must be hard-wired into the ONS and wider statistical system culture, in order to cope with the challenges of a changing world, such as digitalisation and AI. The response to the COVID-19 pandemic shows this can be done.

5.1.2 Develop surveys that are easy, and possibly automatic, to participate in, ensuring that the value of participation is well-communicated to the survey respondents.

Data

5.1.3 **Continue progress in making use of new data sources to mitigate the adverse impact of financial and other resource limitations. Link with business, society, charities, and academia for both data and insight.**

5.1.4 **Explore direct access to business data for statistics producers.**

5.1.5 **Improve capacity for sharing and linking data. Mandate government departments to share data in full and work with them to make this process as straightforward as possible.**

5.1.6 **Improve local level data to make official statistics more representative and relevant.**

5.1.7 Continue developing the High Street Data Service and Partnership to collect information about the health of high streets at a local level.

New and improved statistics

5.1.8 **Identify key and vital economic statistics to prioritise, particularly when resources are constrained, with the aim of improving trustworthiness and reliability.**

5.1.9 Changing work patterns and atypical business models need to be tracked.

5.1.10 Track unpaid work, care givers, the relative burden on and experiences of women and men, and the interaction with the labour market.

5.1.11 Enhance self-employment statistics.

5.1.12 Update job categorisation in the context of changes due to work roles due to the impact of AI.

5.1.13 Produce economic statistics that are more relevant to individuals, taking account of household and geographic disparities. For example, the trend of GDP may be in stark contrast with citizens' experience of stagnating wages.

5.1.14 Revive the e-commerce survey to include modules on AI use, cloud computing, software development and robotics.

5.1.15 Engage with other national statistical organisations and academics to ensure international comparisons and to benefit from the developments in other agencies.

5.2 Coherence

> What are the priority topics that need to be coherent between the four UK nations?

Needs

5.2.1 Coherence and harmonisation work requires dedicated funding, in the context of very disparate resources in the various statistical agencies.

5.2.2 **Establish a forum to develop a plan for better coherence across the statistical system.**

5.2.3 **Coherence is required at multiple levels, including internationally as well as across the four nations, regions, local areas, and across topics. This requires collaboration and perhaps balance, as well as guidance and clarity on whose role it is to ensure coherence.**

5.2.4 Coherence needs to be kept in mind as data sources evolve.

Opportunities

5.2.5 Engage with users across the UK to identify and address easily resolved issues in coherence, such as harmonising weekly death figures.

5.2.6 Consider if it is possible to unpick existing measures from different devolved administrations to produce something 'comparable enough'.

Challenges

5.2.7 Administrative data pose particular challenges of coherence, and methodological developments are needed, with a priority to share expertise across the different agencies.

5.2.8 Recognise that some sectors, such as health, education and the environment face particular challenges in achieving coherence, requiring advocacy and collaboration among ministers, policymakers, data experts and users.

5.2.9 A balance needs to be achieved between locally specific data, relevant to local needs, and harmonised data to inform comparisons and allow aggregation across areas. This may become more difficult as devolution is increased even within the nation states.

5.3 Data sourcing, quality and methods

> Do we need an overall strategy for data sourcing to drive quality and resilience and ensure transparency to users?

Administrative data

5.3.1 **More use must be made of administrative data, linked survey data and the use of the former to improve the latter. This requires coordination, including with non-government organisations collecting data.**

5.3.2 Official statisticians need to have more influence on the design of digital data architectures for administrative data.

5.3.3 Closer engagement with data collection and administrative data structures will improve data quality and hence reduce the work in data cleaning.

5.3.4 There is a lack of coordination and interoperability in the data landscape which needs to be tackled.

5.3.5 Use administrative data to improve small area statistics.

5.3.6 **More use of sources of data external to government.**

5.3.7 Develop and publish models for combining survey and administrative data to produce integrated statistical outputs.

Granularity

5.3.8 More timely, inclusive, and granular statistics are a priority. Monitor closely the merger of local authorities, since this could lead to loss of granularity.

Quality

5.3.9 **Involve statisticians at the design and data collection stage for administrative data.**

5.3.10 Improve data literacy among frontline workers to improve data quality, and understand how to help them get value from the data.

5.3.11 Develop and communicate appropriate metrics of data quality.

5.3.12 Maintenance of existing statistics is important, not merely producing statistics on new topics.

5.3.13 There are increasing costs of maintaining data quality. Budgets need to be considered. Is a quality deterioration sometimes acceptable for some statistics and how would such a decision be made?

5.3.14 Consistent identifiers could alleviate the very considerable time analysts spent cleaning and merging data and could improve the quality of merged data.

5.3.15 There are a range of views around ID numbers (opposition as well as support), but linkage without ID numbers is expensive and time-consuming so there are significant costs in not having such a system. A strategy for public buy-in is needed.

5.3.16 More guidance on handling discontinuities in time series arising from changing data collection methods is needed.

Recruitment

5.3.17 The challenge of recruiting statisticians, given the competing opportunities needs examination. Skills need investment. The GSS should review the training opportunities, consider alternative career pathways and look for 'out of the box' exceptional ideas on boosting the people pipeline, with external partners.

5.4 User engagement

> How should the statistics system enhance user engagement and communications to better ensure user needs are met?

Accountability, transparency, and process

5.4.1 Heads of profession need to identify good and bad areas of user engagement.

5.4.2 **All online published government statistics should include a 'Comments' box, permitting user feedback and better understanding of who the users are.**

5.4.3 Promote awareness that user engagement is a key part of producers' job, not an optional extra.

5.4.4 Increase transparency through development plans and maintenance of a live document of unmet needs. This could also reduce the need for repeated enquiries.

5.4.5 User engagement should be a feedback loop, not a one step activity in either direction.

5.4.6 **Producers should have a mechanism to explain their decisions and prioritisation, and why things are not happening. This can point out the need to balance competing demands. Representatives of the public should be engaged in resource prioritisation.**

5.4.7 Ensure that all user enquiries receive timely and satisfactory responses, and that email inboxes are monitored regularly.

5.4.8 Ensure transparency in investigations of errors in the published statistics.

Rationalisation of engagement strategy

5.4.9 **Develop a joined-up strategy formalising the approach to user engagement that considers balancing views, explaining decisions and supporting theme groups. One might think of this as 'stakeholder relationship management'.**

5.4.10 Provide more support to user groups, including resources and assistance from new members of the GSS. Consider partnerships in this work.

5.4.11 Recognise cultural barriers to user engagement, especially for under-represented communities. Perhaps involve a specialist communication team – see Further Points below.

5.4.12 Engage with Connected by Data, which is running a community of practice for public servants who are trying to engage the public.

Consultations

5.4.13 Consultations should not merely cover data collection, but also outputs and dissemination. It was noted that ONS consultations already do, but other relevant bodies may not.

5.4.14 The Code of Practice for Statistics should clarify the definition of what is and is not a consultation.

Further points

5.4.15 Local authorities do not feel sufficiently engaged. Second or co-opt local government staff into GSS. This is a general point which might be considered with various classes of users. Incorporate the principle of co-design in consultations, especially where local authorities are providing data.

5.4.16 Archive past records of StatsUserNet before its relaunch. [Note: a solution to achieving this has been found since the Assembly.]

5.4.17 Are statisticians the right people to lead on user engagement? Perhaps better to have a team of communicators, with statisticians feeding into that.

5.4.18 Publish good stories on data usage.

5.4.19 User engagement involves resources.

5.5 Health disparities

> How should we improve measuring health disparities across different communities, including health surveillance and monitoring?

Data

5.5.1 The data for under-represented groups is often of poor quality.

5.5.2 **Need to improve the quality of ethnic coding and equality, diversity and inclusion (EDI) in general in health records. Current practices often use inadequate categories and updated guidance is needed for accurate self-reporting. Implement modernised census ethnicity codes in NHS data sets, transitioning from the 2001 to the 2021 Census codes.**

5.5.3 There was a call to collect more granular data, especially in terms of geography, to ensure better data interoperability and insights. However, granularity is often traded against cost and quality, so some sort of best compromise should be aimed for.

5.5.4 **Need for considerations of coordination for UK-wide data, particularly with the significant devolution in the health and care system.**

5.5.5 There are significant potential benefits from using more third sector data and the wider health data ecosystem.

Methodology

5.5.6 The need to understand how individual characteristics, and combinations of characteristics, impact health disparities.

5.5.7 The GSS should examine the relationship between factors like poverty, deprivation and specific health outcomes (for example, lung conditions).

5.5.8 Develop a method to provide up-to-date population denominators by ethnic group and deprivation for accurate health inequality analysis.

5.5.9 Improve trustworthy data linkage, both locally and nationally, and combined survey and administrative data.

5.5.10 Improve access to linked data sets for local authority public health teams and their public health intelligence teams.

Communication

5.5.11 **Improve the important communication strand explaining to citizens how their data is used in making better decisions.**

5.5.12 There is some concern that a shift to administrative data might lead to loss of critical insights. Ensure the continuation and support of surveys like the Health Survey for England to maintain comprehensive data collection.

5.5.13 Develop and disseminate an easy-to-understand privacy notice to improve public trust and awareness of the value of data collection and usage.

5.5.14 Review and potentially adjust current data publication practices to enhance access to valuable data, reducing barriers for public health teams.

5.6 Public sector performance

> What's needed to assess the performance of public services, such as the health service?

Evaluation

5.6.1 **There can be a tendency to assess performance by focusing on efficiency, but effectiveness and outcomes are also obviously critically important aspects of public service performance and so need to be measured.**

5.6.2 **Develop performance metrics and plans in advance so that effectiveness of policies can be determined.**

5.6.3 Performance measures should be developed by a consensus of experts, not politically driven.

5.6.4 Make more use of properly designed experimental approaches to test and evaluate public sector strategies and decisions.

5.6.5 **Be aware of the gap between high level statistical descriptions and people's lived experience, since this can harm public trust. Producers should try to ensure statistics have relevance to lived experience where possible.**

5.6.6 This is misalignment between statistical geographies and where service level delivery happens.

5.6.7 Make public service performance data more accessible and transparent to the public and media.

5.6.8 Increase transparency in tracking NHS financial data to ensure clear understanding of budget allocations and expenditures.

Data

5.6.9 Need high-quality data on people's experiences to improve services. Engage with bodies such as Healthwatch England on this.

5.6.10 Where there are different or contrasting approaches in different nations and regions, these can be used to learn which is the most effective. Enforcing coherence eliminates this potential.

5.6.11 Clarify the question being asked, so that the right data to answer it can be collected, rather than making do with what is there.

5.6.12 There can be a noticeable trade-off between quality of data in terms of timeliness and accuracy.

5.6.13 Improve data sharing across services to better track the movement of individuals and identify heavy users of services.

5.6.14 Standardise the approach of data collection and publication across different departments to improve quality and usability.

5.6.15 Need harmonisation of what we already have and of what data is being collected.

Other points

5.6.16 Evaluate and potentially reduce the reporting burden on frontline services to avoid diluting focus and overburdening staff. Make it easier to collect data.

5.7 Data linkage

> What are the levers needed to enable better data sharing and linkage? How should the statistics system communicate with and involve the public on the ethical issues around the extent of data sharing and linking?

Public acceptance

5.7.1 **Enhance communication efforts to clearly articulate the direct benefits of data linkage and sharing to the public.**

5.7.2 Importance of transparency and safety for public trust.

5.7.3 Questions about consent, safety and transparency of administrative data are often raised, but is the linkage of survey data always consensual, safe, and transparent?

5.7.4 Find ways for statistical organisations to demonstrate trustworthiness by showing an understanding of public views.

5.7.5 Public collaboration and alignment of standards are essential, including on the specific issue of the value and acceptability of identifiers.

5.7.6 Develop a national public engagement programme to gather public input on data sharing and linkage policies. We are aware of work on this by NHS England, the former Administrative Data Research Network, and others, but the fact that this point was made suggested awareness of the exercises needs to be raised.

5.7.7 It is necessary to find ways to communicate research outcomes in ways that non-researchers, including policymakers and MPs, can understand.

5.7.8 Explore the potential for more local decision making and accountability in data sharing and linkage processes.

5.7.9 Develop strategies to address public concerns about data being sold and ensure clear communication about the ethical use of data.

Data

5.7.10 A standardised approach to data handling is desirable.

5.7.11 Consider how the data for people with specific privacy concerns, such as refugees from authoritarian regimes or victims of domestic violence, is used.

5.7.12 Make use of privacy enhancing data analysis technologies.

5.7.13 Explore whether every department should prepare their data in a form suitable for the Integrated Data Service, so that the GSS can show real value faster. Moreover, the National Data Library could be not only be a library of data, but also of agreements, encouraging transparency for the public and for organisations requesting data in future.

5.7.14 Research to what extent the opportunities for sharing and linking are reduced (or made more difficult) by outsourcing data collection, and how this might be mitigated.

Governance

5.7.15 **Fragmented and variable approaches to data governance and safeguards across different data controllers make both data linkage and the crucial task of explaining to citizens how their data is used, more difficult. A more streamlined and coordinated approach to data access governance and linkage across data sets could drive forward research and facilitate better public transparency and understanding.**

5.7.16 The ONS or GSS should establish a ‘trust centre’ on the website to provide transparency about data usage and security protocols, similar to the models used by Stats Canada and CSO Ireland.

5.7.17 **The ONS or GSS to develop standardisation in both soft procedures and technical infrastructure. Public think of government as one entity, but different agencies (local and national) have varying infrastructures, and coherence between them is essential. Create a standardised governance framework for data handling that includes common language and principles, and ensure all trusted research environments explain their position against these principles.**

5.7.18 Clarity is needed on subsequent data linkage use by organisations that do not actually own the data.

5.8 Net zero

> How can we best bring net zero-related official statistics together to tell a coherent and complete story at national and local level?

Communication

5.8.1 Emphasise story and communication. Different ways of expressing things are appropriate for different places and people but we need a coherent overall narrative.

5.8.2 The importance of understanding how people respond to and engage with information. What are the drivers of citizen behaviour?

5.8.3 Part of the challenge of building a coherent narrative is that one can appear to be making progress on one measure while going backwards on another.

5.8.4 **Create a centralised, cross-governmental dashboard to communicate net zero progress to the public, like the COVID-19 dashboard.**

5.8.5 **Address concerns about the transparency of revisions in greenhouse gas emission statistics. Generally provide clearer explanations of the changes and methodologies used in climate change and net zero statistics.**

5.8.6 Enhance user engagement by identifying key questions and data needs from various stakeholders and incorporating their feedback into the development of net zero statistics.

5.8.7 Develop the GSS Environment, Climate, and Nature theme to include user forums and events, bringing together producers and users of statistics relating to the theme, and helping build a diverse user community for the theme and a tailored calendar of engagement activity.

Data

5.8.8 There are discrepancies in the data. For example, measurements that the UK benchmarks against are territorial emissions, but this does not account for consumption-based emissions. Territorial emissions are published with a one-year lag and consumption with a three-year lag.

5.8.9 Measurement does not line up with the behaviour that is wanted by policy makers. They want greater use of electric cars, but statistics measure road traffic in general. Need to measure things in a way that aligns to policy-makers' desired change.

5.8.10 The data ecosystem for achieving net zero is complex so that inevitably the statistical system is complex.

5.8.11 Draft an action plan addressing data issues and share it with stakeholders. This plan should in particular cover the data gaps identified in the Climate Change Committee's progress monitoring.

Methodology

5.8.12 Engage with international statistical organisations to develop standardised methodologies for measuring consumption-based emissions.

5.8.13 Develop a scalable data collection infrastructure for local authorities to gather comprehensive data on net zero projects.

5.8.14 Review and improve the methodology for calculating local authority transport emissions to better align with desired behaviours and outcomes.

5.9 The future of the census

> What are the essential quality considerations for the future population and migration statistics system to deliver sustainable, timely statistics about the population?

Data

5.9.1 Recognition that administrative data have considerable potential, but also limitations. Census, survey data and administrative data can work in tandem to identify and resolve drift, bias, and other distortions in data sets

5.9.2 **Concerns about the adequacy of administrative data in this role. In particular some users were anxious about how analysis about households would be possible given that administrative records tend to be based on individuals. Concern was also expressed about whether users of micro-data would be disadvantaged in the future. It was pointed out that people in the margins of society, of particular interest for public policy, are omitted from administrative records. More discussions with users needed. Develop a strategy for public engagement on this matter.**

5.9.3 Not an 'either, or' (census or admin) decision. Some felt that producers were saying they don't need or want a census and whereas users were saying they do need a census, things were more nuanced than this. It is more a question of how to best use the wide variety of data that is available, and how best to collect new data to answer our questions.

5.9.4 Also need confidence in long-term viability and funding of survey data and strategies to cope with discontinuities when data collection methods change.

5.9.5 Importance of small-area data.

Costs

5.9.6 What is the cost for the public sector to fill data gaps if there is no census?

5.9.7 What are the costs and benefits of maintaining live consistent population and address registers? What data sources might feed into this?

Quality

5.9.8 The Census and administrative data are both needed in order to monitor complementary forms of drift and bias.

5.9.9 **Continuity and consistency over time is critical. Need confidence in long-term supply of administrative data. Caution about relying on data availability from the private sector.**

5.9.10 Challenge of discontinuities when data collection method changes.

5.9.11 Issue of different nations running census at different times. Scotland ran its census in 2022 (because of the pandemic).

5.9.12 Address the challenge of distinguishing between multiple households at the same address in administrative data to ensure accurate household statistics.

5.9.13 What are the costs and benefits of developing a mandatory tracking system for internal migration?

Communication

5.9.14 **Need more communication (two-way) of the role, properties, and capabilities of administrative data with users and more broadly with the public.**

5.9.15 This engagement should take note of the value of using flexible table builder to bring forward publication of data.

5.9.16 Need for continued consultation. Conduct a consultation with users to understand their needs and preferences for the 2031 Census, including the potential use of administrative data.

5.9.17 Publish detailed information on the use of administrative data in the last census, including areas where response rates were lower and how admin data was used to fill gaps.

5.9.18 Publish criteria to explain how administrative data will be assessed and how it will meet the standards for accredited official statistics. [NOTE: the ONS has published these criteria since the Assembly.]

5.10 Business statistics

> What data gaps are there in business statistics, for example representation of ethnic minorities, small businesses and more?

Data quality and extent

5.10.1 **Issues with, or missing from, official statistics data are: definition of market; measurement of innovation; metadata on mergers and acquisitions and firm restructuring; linked employer-employee data; the digital economy is not represented in product or industry; longitudinal data collection on small businesses.**

5.10.2 **Standard Industrial Classification (SIC) lacks granularity in the employment in the service sector – outdated statistics. In general, SIC is not representative of the current economy and industry structure and is out of date (we are aware that attempts have been made to update it, and have encountered resistance). But the almost contrary point was also made, that classifications need consistency over time. This clearly needs to be resolved and the GSS needs to consider how.**

5.10.3 Collaboration between statistical system and business is central to improvement.

5.10.4 Lack of interregional data (England), lack of national data across the four nations of the UK.

5.10.5 More data on services needed.

5.10.6 City data important for industrial strategy.

5.10.7 Value of linked employer-employee data sets.

Data collection

5.10.8 Need better methods to collect data, for example by using administrative or more modern direct data collection from business systems.

5.10.9 There is pronounced enthusiasm for administrative data in the business community.

Data availability

5.10.10 Data needs to be more open-source and needs to be accessible by businesses (under secure and safe anonymised environment).

5.10.11 The ONS website needs to improve searchability, rather than presentability or visualisations.

Communication

5.10.12 The Authority's relationship with business statistics producers and users is not strong enough. The Authority needs to engage with the main business representative bodies – for small and large businesses –and perhaps start with no preconceptions. We recommend some research with businesses first, on issues with production and usage, possible data sources, and so on.

5.10.13 Need to explain more on why businesses are being asked to provide data. To communicate the importance of this, more is needed than merely saying "for statistical analysis". It is necessary to explain why it is beneficial for them to provide this data. How can the collection of data from businesses be simplified. Automatic data collection?

Possible savings

5.10.14 Has there been a recent review of which data are essential?

5.10.15 Transition from reporting average weekly earnings to average monthly earnings to better reflect current business practices.

5.11 Equalities

> What statistics are needed to monitor equalities, in particular ethnicity? What data collection mechanisms are needed to better address representation and diversity?

Methodology

5.11.1 **There is a lack of consistency in the 'definitions'. Participate in a review and development of the ethnicity classification as part of the GSS harmonisation function.**

5.11.2 There is a lack of consistency in the 'recording' of ethnicity, sometimes mixing racial, national, and linguistic groups. The same person could end up with multiple identities. Provide clear guidance.

5.11.3 **Too much reliance on snapshot data – need comprehensive longitudinal data. And need to get to a place where individuals are comfortable with their data being used for longitudinal multifaceted studies.**

5.11.4 **Participants stressed the need for clarity in the purpose of collecting ethnicity data. They argued that without a clear understanding of why the data is being collected, efforts to improve data quality and use will be ineffective.**

5.11.5 Increase research capacity within the black and other communities to analyse existing data and provide evidence for social change.

5.11.6 Improved guidelines and coherence around how to describe family types would be useful. Example of difficulty determining whether to say "mixed-sex couple" or "opposite-sex couple", with different branches of government suggesting different choices.

5.11.7 Are there particular challenges of survey non-response bias with these data?

5.11.8 Weigh up the relative merits of disclosure risk versus the value and purpose of publishing detailed breakdowns.

Communication

5.11.9 Data collection may be hampered by a distrust of what the data might be used for. Explain what the data are used for. However, explanation needs to cover more than highly specific potential uses since that would limit the value of the data by restricting its use. Communication with the general population is needed.

5.11.10 More communication with the users needed to ensure the categorisations are fit for purpose.

5.11.11 Data systems are one-way processes. People give the data but do not see it and cannot check it. Facilitate confidence in the data.

5.11.12 Ensure and communicate clear firewalls between operational and statistical uses of data.

5.12 AI and technology

> What are the choices that we should make about the ways in which the UK official statistics system adopts (or does not adopt) AI and other emerging technologies over the coming years?

Role of AI

5.12.1 The official statistics system must embrace AI for the public good.

5.12.2 AI is already proving beneficial, offering automation, efficiency savings, reproducible analytical pipelines, code checking, report production, and sourcing new data types. AI can address real-world issues, such as preventing modern slavery and aiding in prosecutions through techniques like satellite imagery, linking criminal case files, and social media data scraping.

5.12.3 AI can enhance the entire analytical pipeline, from data collection to communication. Avoiding AI would leave the official statistics community behind. To stay relevant, we must work in partnership across the statistics, data science, and AI communities.

5.12.4 AI and novel technologies will play a crucial role in statistics production and communication but we must treat them with caution and scepticism; they are just algorithms.

5.12.5 Statisticians have a crucial role in adopting AI, ensuring statistical rigor, and adhering to the principles of the statistics Code of Practice (CoP).

5.12.6 There was a suggestion that the statistics system should explore working with the tech companies producing large-language models to see how to get the models to use reliable official data in their outputs, but others were sceptical about the feasibility of this.

Trust in AI

5.12.7 **Trust in official statistics must be maintained through transparency, human accountability, and high-quality evaluation of AI models. Transparency requires AI models to be open to scrutiny and code to be publicly available. The CoP's trustworthiness, quality, and value must be adhered to, and the OSR should oversee this.**

5.12.8 Monitoring the impact of AI on public trust in official statistics is essential.

5.12.9 Ethical use of AI is paramount, requiring evaluation for bias to avoid disadvantaging marginalised groups. A diverse workforce and multi-disciplinary teams, including ethics professionals, are key to the ethical adoption of AI.

5.12.10 The issue of consent in the use of AI tools must be considered.

5.12.11 **A coherent approach to validating AI models needs to be put in place, and this needs to be done by independent bodies, not the AI developers.**

5.12.12 There is an opportunity to learn from other national statistical institutes.

Communication

5.12.13 It is very important to communicate effectively with policymakers, decision-makers, service deliverers, and the wider public about the roles of AI and to consider public consultations in official statistics for contentious issues.

5.12.14 The question was raised of whether it matters that users might get statistical information summarised by an AI agent rather than straight from the horse's mouth (for example, the ONS website). This has the merits of efficiency-saving and tailoring communication to questioner's style. However, there is an issue of whether the AI agent introduces misunderstanding or hallucinations. While in principle any AI output must be delivered via a human agent, requiring this would defeat the objective of potential efficiency gain from using AI directly.

Other

5.12.15 Is there a role for synthetic data and digital twins?

5.13 Crime

> Is there anything a single crime measure could add to better inform decision-makers and the public?

Single metrics

5.13.1 A new single measure could go further than existing measures by developing a "crime impact index", based on the impact on or severity for victims, families, neighbours and other groups by comparing the Crime Harm Index, Crime Severity Score (based on effects of crime on society (sentencing), or the Home Office Cost of Crime measures.

5.13.2 A single measure could help attract media attention, galvanise public support, and engage policymakers.

5.13.3 The analogy with inflation indices was made, with the potential for a crime index based on a "basket" of crimes. However, others pointed out that there are several key differences between economic indicators like inflation and crime, and suggested that it wasn't clear that the resource that would need to be invested into it could not be better deployed meeting existing unmet needs of users.

5.13.4 Another analogy was with the Index of Multiple Deprivation.

5.13.5 Having a single index of crime doesn't stop it being broken down for different subgroups. It helps give a consensus view on whether crime is improving.

Beyond single metrics

5.13.6 Single summary measures can be counterproductive for prevention for a variety of reasons (see below).

5.13.7 A single measure fails in that crime is not a single thing. It covers a vast range from trivial crimes to murder. Combining into a single measure seems to fail the public understanding test (how many stolen Mars bars equal a murder?).

5.13.8 We should beware of managing to a single number, which could lead to focusing on specific crimes while neglecting others.

5.13.9 One complexity is that there is differential impact across income ranges – people will experience crime differently based on whether they can afford insurance or replace stolen goods.

5.13.10 Different areas have different crime issues, so it's not obvious that a number computed in one way would be useful.

5.13.11 Improved measures might use weights (for example, weight vehicle crimes by the number of cars or vehicle-owning households) representing the exposed population. For example, youth violence is relevant only to young people, and so should be weighted for that population.

5.13.12 **Develop a crime dashboard. Canada has a good dashboard around victim-weighting figures. People can understand a sentence such as "crime is reducing but crimes that are happening are more severe".**

Incoherences

5.13.13 The biggest issue is marrying the Crime Survey of England and Wales (CSEW) with reported crimes in the context of people's perceptions. The methodological challenge of combining data from administrative sources with that from surveys, when they do not measure exactly the same things, is not unique to crime (as an example it also occurs for unemployment data) and it would be good to see a methodological programme addressing this.

5.13.14 **The CSEW is becoming at odds with what people feel. An abstract measure that gets more removed from lived experiences risks weakening the relationship between measures and perception and can result in a lowering of trust in statistics.**

5.13.15 Beware the risk of feedback distorting the figures (Goodhart's Law) as police focus attention on the measures.

5.13.16 A problem of distorted and biased data also exists. For example, people not reporting crimes for fear of the impact it has on house sales or rentals.

5.14 Labour market

> How can labour market statistics measure the future labour market by best using surveys, administrative data and other data sources?

Labour Force Survey (LFS)

5.14.1 The discussion of LFS issues has not always been helpful. Need to be clear on long-term issues and where we are now.

5.14.2 One delegate suggested that we should not overstate the "flying blind" narrative, pointing out that, although there are legitimate concerns about the quality of the LFS because of falling response rates, it is important to appreciate that this is just one source of data on the labour force and there is a full suite of statistics.

5.14.3 There is need for clarity on the future strategy for labour market statistics, with clear plans published, as soon as possible. Development of these plans should involve consultation with users.

5.14.4 Need to be aware of the challenges of combining administrative data with LFS – there are many differences in the data collected, the populations covered, the definitions used. Need to understand the nuances to combine these data effectively, as relying on only one source can mean we cannot cover some topics (for example, hours worked).

5.14.5 It would be helpful if the LFS could be disaggregated at the Local Authority level.

5.14.6 Explore modular and shorter survey methods to make it easier for respondents to complete the LFS.

Data

5.14.7 Need to continue to increase availability of administrative data.

5.14.8 It is good that the ONS has built up alternative data sources in recent years. However, administrative data cannot replace surveys since some topics require surveys for their collection (for example, job satisfaction, health conditions, and so on).

5.14.9 **Suggest more proactive research on what influences response to surveys and what is an acceptable level of response. This could address what helps – incentives, technology, communicating public duty?**

5.14.10 **Can we completely change the model for survey response: What works for "me"? What works for younger adults, such as online or mobile phone methods? What works for older people typically with less tech access?**

5.14.11 Enhance the LFS to better capture data on unpaid care activities and their impact on economic inactivity.

Communication

5.14.12 **Communicate the next steps for the Transformed Labour Force Survey in the spring update.**

5.15 Local and regional data

> **How can national data and statistics benefit local and combined authorities, and local data benefit national policymaking? What does the statistics system need to do to support organisations in achieving this?**

Impact

5.15.1 Huge potential as biggest impact on public's lives is often through local government. Local is where impact and implementation happen.

5.15.2 **Local data can reach parts national data struggle with: the example of using local data to understand the link between poor quality housing and health in a project to identify homes with damp and mould and their impact on residents' respiratory conditions.**

5.15.3 We should seriously consider whether English regional data is required … [given that] few if any organisations with power to make change work at the regional level.

5.15.4 Is there a need to think about prioritisation and efficiency, for example if the sub national population projections are not useful to local authorities why are they produced?

Communication

5.15.5 The need to reflect people's lived experiences.

5.15.6 **Important to get users together in order to listen to what they need from what's already produced and to understand why they need it, but also to identify data gaps. Co-production is desirable so it is important to involve a range of data providers.**

5.15.7 Promote collaboration among local government organisations to share resources and develop common data models.

5.15.8 Mechanisms for engagement with local councils would be helpful for national bodies.

Data

5.15.9 It is vital that local data remains available. The example was given of a regional charity that contains some of the most and least deprived areas in the whole country. Regional data does not provide sufficient information to allow analysis of the variation within the region. Such data is essential to plan and deliver services.

5.15.10 The issue of providing sufficient granularity needs to be addressed in the national data frameworks, paying particular attention to rural areas.

5.15.11 There is a requirement for significant intersectional data with granular geographies to be able to make decisions and interventions at the local level, though there is an acknowledgement that this increased the risks of disclosure and so on.

5.15.12 The barriers include skills, the availability of resources, competing priorities and the tensions between the needs of local versus central government.

5.15.13 Collaboration could be fruitful with higher education institutes seeking to have an impact. Similarly private and public collaborations could be useful.

5.15.14 National data could be supplemented with local data to capture more detailed insights. This might include running resident surveys, crime surveys, and commissioning bespoke research to address specific issues.

5.15.15 It is important to ensure all national data is available at least at the local authority level and in accessible formats.

5.15.16 A problem exists due to the lack of data resources and skills within rural authorities.

5.15.17 Data sharing issues need to be addressed, particularly with Land Registry data, to facilitate local decision-making.

5.15.18 Connectivity between ONS data and NHS health data is poor. Due to the population rebasing, there has been no data release on survival for over a year (the most recently available data is for 2021), and the release of incidence data has been changed to regional level only and is also behind schedule.

6. International perspective on the UK statistical system

Beyond the sessions on particular topics, the Assembly also hosted presentations from people outside the UK statistical system, to give an international perspective. Key messages from those presentations included that:

- the UK system was highly regarded internationally, and that weaknesses in a particular area should not be taken as representative of the whole
- some issues are wider than the UK system, and have to be tackled internationally and collaboratively
- the importance and challenges of data sharing are universal
- it is important to appreciate that the data infrastructure is greater than merely the ONS
- the breadth of the range of technologies and data sources needed
- the need to show citizens how data was being used was true internationally
- the concept of hardwiring agility and proactivity into the statistical system was enthusiastically recognised
- the need to do more with less resource by identifying lower-value activities, better prioritisation, and taking advantage of technological advance was universal

7. Recommendations for the conduct of future assemblies

We stress that this was the first of a planned series of assemblies. As such, it was partly experimental, with the intention being to learn from it to improve the conduct and content of future meetings.

This first Assembly involved more than merely the day itself. The call for contributions provided a firm basis for preparing an agenda for the day, already demonstrating the value of user engagement. This approach is worth repeating for future assemblies, perhaps also allowing time for feedback from prospective attendees and reflection on issues and topics that do not appear in responses to the call. Moreover, in the days following the Assembly we received a number of email comments from delegates as they reflected more deeply on the discussions. There needs to be some formalisation of this source of more considered thoughts following future assemblies. GSS theme groups could play a major role in this engagement.

Inevitably, with a limited amount of time on the day itself, a choice had to be made on what topics to cover in the breakout sessions. While future meetings might return to all or some of these topics, other topics must also be considered, and future calls should invite suggestions. Other suggestions which have already been made are consumer behaviour or attitudinal data, cost of living, education, higher education, transport, food and agriculture, defence, the value of following international standards (especially in the context of Brexit), and a session on how to create an official statistical system with more porous boundaries so that expertise from outside can be utilised (such as via secondments). It might be appropriate to include a session where attendees look at suggestions and prioritise proposals (perhaps in a second day) since the current approach made it difficult for the Authority to get a view on what delegates themselves would prioritise.

It would be helpful, and would encourage future participation and engagement, if details were given of how the outputs from this Assembly will feed in to the Authority's work over the next few years.

Further recommendations for future assemblies included:

1. Enable other ways for people to input in person on the day, as it was not always possible to feed in to sessions (for example if one wanted to participate in two parallel sessions or had clashing meetings).
2. Perhaps enable online breakout rooms for the online participants, to take place over the breaks.
3. Facilitate greater representation from policy professionals, whose perspective was missed.

4. Encourage greater participation from sectors which might benefit from greater representation, such as business and local government.
5. At future assemblies, all delegates should be able to see others' responses, whether in the notes taken at the meeting or via Slido, and these should be generally available after the meeting.
6. It might be helpful to have a session focused on what will be removed from the Authority's workplan. The Authority could lead this session, for example providing a list of questions (for example, should we do X, or dump Y, devote more or less resource to Z?). This would bring home the practicalities arising from resource limitations.
7. Some sort of cost-benefit analysis of the Assembly should be made. This might be in terms of changes which occurred as a consequence of the Assembly.
8. The Authority should identify which of the concerns raised by users are already a focus of attention, and feedback such a list to NSEUAC.
9. Perhaps an online survey about what delegates see as priorities could be conducted. It would be interesting to see a before- and after-Assembly comparison of the results.
10. The Assembly delivery group created a framework for the sessions, but not all sessions adhered to that framework, perhaps because alternative approaches best suited the topic. Nonetheless a standard framework at some level does facilitate capturing user concerns.
11. An important issue raised was how often the Statistical Assembly should be held. The Lievesley Review said triennially. However, there is some enthusiasm for more frequently, not least because of a feeling that people want the statistical system to be more nimble, agile and responsive. Annually seems too often, given the number of people involved and the resources and effort required, so perhaps biennial would be suitable.
12. An elaboration of the biennial suggestion was to run every two years, with a full Assembly like this first alternating with a more focused meeting (so that the full Assembly took place every four years). The focused meeting could be devoted to particular topics, or could be jointly organised with other bodies (for example, the RSS and the Confederation of British Industry).

8. Acknowledgments

The members of the NSEUAC would like to express our appreciation to Professor Denise Lievesley and Jonathan Everett for helping with the preparation of this report. We would like to thank the UK Statistics Authority and the Royal Statistical Society for organising the event, the Assembly delivery group who put together an immensely stimulating meeting, the Secretariat, who laboured behind the scenes to ensure that it all functioned smoothly, and also, of course, the 556 delegates without whose wise thoughts and informed comment none of it would be possible.