UK Census User Conference 2025

Abstracts

Keynote presentation

The importance of user and public voice in official statistics
Jeni Tennison, Founder and Executive Director, Connected by data

Members of the public aren't just users of statistics – they are the people statistics are about, and their lives are affected when statistics are used. What kind of say should they have over statistics and how can statistics producers work with them to ensure public acceptability? In a recent project with the Office for Statistics Regulation, Connected by Data has been exploring public involvement in statistics. In this talk, Jeni will discuss what they found and where the challenges, and opportunities, lie.

Parallel session 1a

The changing geographies of ageing and age-mixing in Scotland
Rachel Wilkie, Nissa Finney, Qiong He, Alice Butler, Jo Mairi Hale, and Elspeth Graham, Centre for Population Change and University of St Andrews

Like other High-Income countries, Scotland is experiencing rapid population ageing, with evidence of spatial polarisation of age groups. This study uses data from the Scottish Censuses of 2011 and 2022 to address questions on intergenerational justice and understand patterns of ageing and age-mixing. With attention to urban-rural differences, we examined the distributions of people over the age of 65 living in each data zone; spatial segregation between older and younger people (Dissimilarity Index); and local age-mixing/diversity (Simpson's Diversity Index).

Our results show clear geographic patterns of change in age diversity between 2011 and 2022: in general, remote and rural data zones experienced decreased age diversity whilst urban data zones experienced increased age diversity, with rural and urban neighbourhoods both seeing an increase in the older population. Notable emerging geographies of ageing are that while most large urban data zones are becoming older and increasingly diverse, a relatively high proportion are experiencing the opposite trend of an increasingly youthful age structure and decreased age diversity. Furthermore, we found that a relatively high proportion of data zones in accessible rural areas, predominantly in the Central Belt, characterised by a younger age composition and increased age diversity.

This paper exploits the unique small-area qualities of Census data to contribute new findings on the spatial and temporal patterns of ageing and age-mixing in Scotland, with relevance for broader debates on health and social care provision, housing accessibility and equity, and population sustainability.

The shape of Britain: Topologically mapping the census
Simon Rudkin, University of Manchester

The topology of a dataset is defined by the joint distribution of all of the variables that comprise the dataset. The UK Census provides a comprehensive snapshot of the socio-demographic make-up of the country. Necessarily, the multitude of variables within the Census mean that the joint distribution of census data is hard to understand. Tools of Topological Data Analysis (TDA) help us to make sense of the wealth of data at the fingertips of researchers. We see how TDA maps the joint distribution. This talk will use specially generated examples to show the topology of the socio-demographic

characteristics of the UK as captured in the Census. Further, consideration of historic census data allows evaluation of the changing construct of the UK.

To demonstrate the power of TDA the examples presented in the talk will be:

- Educational inequality using parliamentary constituencies and wards. Here we will see how the changing aggregation of data changes inference in the data maps
- Changes in dominant employment types using the NSSEC classifications between 2011 and 2021. The example will show how change manifests within the data

The talk will also show how data maps and geographic maps can be combined. Where proximity in geographic space implies similarity in longitude and latitude, proximity in data space implies similarity in all variables. Proximity in data space does not imply proximity in geographic space.

This presentation targets inspiring researchers to use TDA to gain deeper understanding of the UK Census data space and foster impactful collaborations using TDA.

The geographies of religion in England and Wales, 2001-2021
AbdulQadeer Fayaz-Khan, Gemma Catney, Christopher Lloyd and Momoko Nishikido, Queen's University Belfast, and Jesse Ransley, Office for National Statistics

This presentation examines the shifting geographies and demographics of religious communities in England and Wales, tackling three key issues: the evolution of major religious groups over time, the geographical diversification of religious affiliations, and the current state of religious diversity amid secularisation trends. Drawing on Census data on religion for 2001, 2011 and 2021, the study uses the Dissimilarity Index (D) and Reciprocal Diversity Index (RDI) to analyse residential segregation and diversity at the neighbourhood (Lower Layer Super Output Area (LSOA)) and Local Authority levels.

The findings reveal significant demographic changes. The proportion of people describing themselves as Christian - some 46.16% of the England and Wales population - continues to decline, while non-religious identities rise sharply. The proportion of people identifying as Muslim (6.51%) and Hindu (1.74%) now comprise the second and third largest groups.

Analysis of segregation shows decreasing residential separation for people identifying as Muslim, Sikh, and Hindu, reflecting growing integration, while people describing themselves as Jewish maintain relatively high segregation levels, essentially unchanged since 2001. People describing themselves as Christian had low segregation levels in 2001 and these are even lower by 2021.

Ongoing research will show how religious diversity nationally, regionally, and in local areas, has evolved in recent decades. We will explore the growth and spread of religious diversity, and the ways in which this relates to the new neighbourhood geographies of ethnic diversity evidenced in allied research (Catney et al., 2023).

Parallel session 1b

Introducing the Socioeconomic Index for Small Areas (SEISA): A new UK-wide areabased measure of deprivation
Tej Nathwani, Siobhan Donnelly, and Archie Bye, Higher Education Statistics Agency

The Indices of Deprivation represent the most commonly used area-based measure of deprivation in the four nations of the UK. Yet, with each country adopting its own methodological approach in creating the Indices, this measure lacks UK-wide comparability. The size of the areas used in forming the Indices can also sometimes lead to it being difficult to identify pockets of deprivation within less deprived neighbourhoods. Furthermore, the final multiple deprivation index in each nation is known to be less effective in capturing deprivation within rural areas too.

This presentation will introduce the Socioeconomic Index for Small Areas (SEISA) – a new UK-wide area-based measure of deprivation formed initially using Census 2011 data and now updated using Census 2021/2022. The methodology involves using smaller-sized areas (in terms of population size) than those used to produce the Indices of Deprivation to reduce the possibility of not picking up deprived localities that are surrounded by less deprived vicinities. SEISA has been formulated using information relating to the qualifications and occupations of residents in these smaller areas, which emerged in our literature review as key determinants of whether a household/individual may have a low level of income.

We will discuss the work we carried out to validate our measure captured deprivation as anticipated, as well as the consistency of the findings across the two Census collections. We shall also highlight some of the other advantages SEISA can offer to data users, including its ability to more effectively pick up deprivation in rural areas.

An environmental justice analysis of two decades of air quality change in England in relation to changing deprivation and ethnic group population distributions
Paul Norman and Gordon Mitchell, University of Leeds

There is strong evidence for environmental inequality in the UK, including higher air pollutant concentrations in more deprived areas. Whilst air quality improved between 2001 and 2011 it did so most slowly in more deprived areas which also bore a disproportionate share of declining air quality including non-compliance with statutory air quality standards (Mitchell et al., 2015). With the availability of 2021 Census data, we develop a twenty year picture of the changing social distribution of air quality (NO2, PM2.5) for LSOAs in England. Our social metric is the Townsend deprivation index using a time- and geography-comparable measure (Norman et al., 2024). We introduce an ethnic dimension (White, South Asian, Black, Mixed, Chinese & Other population groups) to further explore how air quality has evolved over the long run.

The results reveal a strong, persistent correlation between poor air quality and high deprivation. A large proportion of the ethnic minority populations experience the double disadvantage of living in areas of high deprivation and poor air quality; most notably the Black groups. Gini coefficients illustrate, however, that for Mixed ethnic, South Asian and Black groups there is a slight easing of relative exposure to poor air quality over time; perhaps due to suburbanisation / dispersal away from urban centres. However, overall compliance with the more demanding World Health Organisation guideline air quality standards is poor for both pollutants and exhibits strong social gradients. Reducing these environmental inequalities, and the health inequalities they imply, will be an ongoing challenge.

Mitchell G, Norman P & Mullin K (2015) Who benefits from environmental policy? An environmental justice analysis of air quality change in Britain, 2001-2011. Environmental Research Letters https://iopscience.iop.org/article/10.1088/1748-9326/10/10/105009

Norman P, Lloyd C, McLennan D, Ferguson S & Catney G (2024) 50-year Deprivation Trajectories: Local Area Change in England and Wales, 1971–2021. Applied Spatial Analysis & Policy https://doi.org/10.1007/s12061-024-09583-w

Trajectories of deprivation in the UK, 1971-2021
Chris Lloyd, Sara Ferguson, and Gemma Catney, Queen's University Belfast, Paul Norman, University of Leeds, and David McLennan, deprivation.org

Deprivation measures are used widely for understanding spatial inequalities, for designing policies, and for targeting resources at vulnerable communities. Only rarely are the implications of the history of deprivation considered directly. Areas may have similar levels of deprivation at one time point, but very different levels a decade before. Thus, strategies for reducing deprivation may be appropriate in one area with high deprivation, but not in another with a very different deprivation history. In this paper, the

focus is on how deprivation levels have changed over a fifty-year period in neighbourhoods across the UK. Census data for 1971 to 2021 (2022 in Scotland) are used to construct a deprivation measure - the Townsend index - using data on employment, housing tenure, car or van access, and overcrowding. Results for individual UK nations show common trends, with an absolute decrease in deprivation in most neighbourhoods but with differences between more deprived and less deprived areas remaining relatively constant. The final part of the paper considers the challenges of producing a UK-wide trajectory classification given differences in data definitions and geographies. It presents provisional results which point to, for example, common experiences in many areas across the nations but with some distinctive differences in terms of spatial concentrations of persistent deprivation. It is argued that the Census offers crucial insights into how deprivation and population and housing characteristics have changed in small areas across the UK and that this information should be used more widely to design schemes to better tackle spatial inequalities.

Parallel session 2a

The ONS Longitudinal Study of England and Wales: The research possibilities of its large sample size and 2021 Census linkage
Alison Sizer and Stephen Jivraj, UCL, and Lucinda Platt, LSE

This paper introduces the Office for National Statistics Longitudinal Study (ONS-LS) and its data. It highlights the upcoming linkage of the 2021 Census data, and the opportunities it offers to examine a range of new topics and socio-economic/ demographic changes taking place since the 2011 Census.

The ONS-LS follows a sample of the England and Wales population from Census data across time (1971-2011), linking in administrative data (births, deaths). ONS-LS members are selected based on four birthdays, providing a 1% representative sample. Individuals enter through birth/ immigration and leave through death/ emigration. Its main strength is its large sample size (>1 million), allowing the analysis of small areas and specific population groups.

Linkage of 2021 Census data will provide opportunities to examine:

- new topics (gender identity, sexuality, veteran status).
- longitudinal transitions related to questions first included in 2011 (national identity, passports held).
- changes occurring during the 2011-21 intercensal period when the UK's EU referendum and subsequent withdrawal occurred.

2021 linkage is currently in its beta testing phase. This phase will test the 2021 data for accuracy, comparability and consistency, and use feedback from test participants to decide whether to release the database to the wider research community. We will present preliminary findings from our own beta test project investigating whether "Brexit" was associated with changes in migration flows of EU nationals to/from England and Wales over the 2011-21 period, and the extent to which the referendum was associated with British citizens accessing EU member passports, to retain the right to free movement.

The longitudinal impossible dataset: Helping users navigate the ONS Longitudinal Study
Andreas Mastrosavvas and Nicola Shelton, University College London

The Office for National Statistics Longitudinal Study (ONS-LS) follows a 1% sample of the population of England and Wales through each decennial Census, linking Census data with data from birth, death, and cancer registers. As one of the largest datasets of its kind in the UK, the ONS-LS is used for public good research on topics ranging from public health to labour market outcomes. However, access to the data is highly controlled and only possible via secure settings, meaning that researchers must often identify required variables and develop code prior to seeing the data. With thousands of variables

available, navigating and exploring the available metadata can be a complex task for users, which also often results in additional administrative burden on user support teams.

This presentation will showcase the Longitudinal Impossible Dataset (LIDS): a customisable artificial dataset intended to familiarise prospective users with the data structures and variable domains represented in the ONS-LS. Areas covered will include conceptualisation, development, deployment, and user feedback, sharing insights for practice in secure data user support services for social science research. It will also offer a brief introduction of related initiatives undertaken at the Centre for Longitudinal Study Information and User Support (CeLSIUS).

Long-run implications of housing ownership reform: Intergenerational evidence from Right-to-Buy
John Gathergood and Yifan Li, University of Nottingham, Richard Disney, University of Sussex, and Stephen Machin and Matteo Sandi, LSE

The Right-to-Buy was one of the world's most radical homeownership policies, responsible for an increase in the national homeownership rate in the UK from approximately 60% in 1979 to over 70% by the early 2000s. The effects of the policy continue to persist today and are a topic of ongoing national debate. This study uses the long-run intergenerational linkage of the census data available in the ONS Longitudinal Study to estimate the effects of right-to-buy on early and mid-life outcomes of children affected by the policy. Exploiting exogenous eligibility to purchase a home based on length of tenure, and using an instrumental variables strategy, we estimate the effect of the policy on long-run outcomes relating to education, employment, family formation, health, and intergenerational transmission of housing tenure. The study draws on the long time span of census data, from 1971 through to 2011 and will make use of 2021 census data in due course. The detailed inter-family linkage information, together with the broad range of outcomes collected in census data, permits a rich analysis of multiple outcomes over multiple time periods. Among the outcomes of interest, the project examines the impact of the policy on the housing tenure of children impacted by the policy when they reach mid-life, as measured in the 2021 census data at approximately age 40. This study will constitute the first comprehensive evaluation of the Right-to-Buy policy on the lived experience of subsequent generations who childhoods were impact by the policy.

Ward typology for community resilience in England
Maria Mercedes Fleitas Delgado and Simon Rudkin, University of Manchester

The understanding of community types is relevant to the analysis of community resilience as it ensures that resilience strategies are context-specific. This research uses 2021 Census data to present a typology of community resilience of England at the ward level, using the Topological Data Analysis Ball Mapper (TDABM) algorithm. The study captures the complexity and interconnectedness of communities by exploring identified community resilience characteristics using the PEOPLES Framework as reference for variable selection. The TDABM approach effectively identifies distinct community profiles and highlights overlaps in community attributes. To assess the consistency and robustness of the classification, k-means clustering is applied as a comparative method. The analysis suggests the need for resilience strategies that recognize the diversity and complexity of community types. Despite data limitations at the ward level, the findings offer valuable insights for policymakers and researchers aiming to enhance community resilience across England. This research contributes a novel application of TDABM and demonstrates the value of TDABM for analysing socio-demographic Census data.

Parallel session 2b

Ethnicity and internal migration in England and Wales
Momoko Nishikido and Gemma Catney, Queen's University Belfast

Despite considerable interest in the residential redistribution of White and minority ethnic groups, fundamental questions remain on who is migrating and where, and on the role of (internal) migration in shaping ethnic geographies. This paper explores patterns of internal migration in England and Wales by ethnic group using 2021 Census data commissioned from ONS. The data provide information on change of address one year prior to Census day by ethnic group and age, at three spatial scales (Districts, and Middle and Lower Layer Super Output Areas).

We first explore the migration propensities of each ethnic group by age, and how these rates have changed 2001-2021. The paper next builds on previous research which has highlighted commonalities in migration away from urban concentrations and towards greater inter-group mixing. Recently-renewed political-policy claims of so-called 'self-segregation' into minority ethnic clusters contradict new evidence from Census data analyses across three decades, showing a steady decline in residential segregation of all ethnic groups alongside a growth in neighbourhood ethnic diversity. We use age-standardised net migration rates to explore ethnic group differences in intra-area residential mobility and migration towards and from own-group ethnic concentrations. Preliminary results from this novel multi-scale analysis suggest net migration for minority ethnic groups away from areas of highest co-ethnic concentrations. A similar, but more complex, pattern at the neighbourhood (LSOA) scale points towards net gain in White majority areas for all ethnic groups, mirroring the growth of ethnic diversity in these spaces between 1991-2021.

Refugee Integration Outcomes (RIO) insights: Data linkage methods and analysis from linked Census 2021 and administrative data for asylum and resettled refugees in England and Wales between 2015 and 2021
Daniel Jones, Home Office

Robust quantitative data on long-term refugee integration outcomes is lacking. Existing studies on refugee integration outcomes often exhibit weaknesses in the robustness of their data and tend to be based on qualitative research. The Refugee Integration Outcomes (RIO) longitudinal cohort study aims to address this evidence gap by creating a comprehensive dataset of refugee outcomes to better inform our understanding of integration and provide a more robust evidence base for interventions by government and other providers.

Home Office refugee data was linked to Census 2021 and administrative datasets including Personal Demographic Service, Home Office Borders and Immigration data, Births and Deaths through deterministic linking, associative matching, and clerical review. The current RIO dataset includes refugees resettled under the Vulnerable Persons or Vulnerable Children's Resettlement Scheme (resettled refugees) and asylum seekers who were granted refugee status between 2015 and 2020 (asylum route refugees).

There was a large difference between resettled refugees and asylum route refugees in the quality of data matching. The difference in linkage rates emphasises that these are two distinct cohorts with different characteristics and behaviours.

We present results from analysis of the linked to Census 2021 data that shows improved economic activity for those who have been in the UK for longer, an association between language proficiency and employment, and the differences in housing outcomes of the two cohorts. We hope that these quantitative insights enable us to understand their integration outcomes longer-term and will be helpful in informing policy to help this vulnerable population integrate into their communities.

Gender inequality in general health among young people: Investigating the contribution of socioeconomic indicators included in the 2021 Census microdata
Katrin Metsis and Frank Sullivan, University of St Andrews, Joanna Inchley, University of Glasgow, and Andrew James Williams, University of Edinburgh

Estimating health inequalities among adolescents and young people is complex because of the measurement of socioeconomic position (SEP) and health. Census data includes several validated measures of SEP and health. General health (GH) is a global summary measure that predicts mortality and morbidity and is a useful measure of health for young people. The National Statistics Socio-economic Classification (NS-SEC) is based on employment relations and conditions; at the level of the household's reference person, it indicates the resources available for young people.

Methods
This study focuses on gender differences in health inequality in 10 to 24-year-olds. We utilised the 2021 Census Safeguarded Individual Microdata Sample at Region Level. Gender-stratified bivariate analysis and logistic regression were applied to examine GH by the NS-SEC of the family reference person, controlled for age and household deprivation of housing and education. We also investigated the effect of disability status; and individual-level NS-SEC and educational level among 19-29-year-olds.

Results
257, 819 boys and 247, 314 girls were included in the analysis. Compared to young people from managerial and professional households all other groups had higher odds of reporting poor health. Young people from households of never worked or long-term unemployed reference person had more than two times higher odds. These findings are in line with our analysis of the 2001, 2011 and 2021 census datasets (Metsis et al. 2024); and support the Fundamental Causes Theory that views SEP as a fundamental cause of health inequalities as people from different backgrounds have different access to health-supporting resources.

Sorting low income workers
Concetta Gigante, University of Liverpool

This paper studies the impact of the Universal Credit programme in England on the intensive margin of labour by using the 2021 UK Census data. The Universal Credit is an important source of income for low income workers in employment. The paper proceeds in two parts. In the first part, I present evidence on low income workers in the UK based on their occupation, deprivation and industry by region. In the second part I investigate the impact of the increase in the benefits of the programme in November 2021 and implement a regression discontinuity design to examine the impact of the threshold of age eligibility on low income individuals and their hours close to the threshold of receiving the benefits.

Session 5: Research paper

Royal British Legion Needs Analysis: Using Census data to understand the needs of the armed forces community and inform design and delivery of support to meet needs
Morgan Brown and Ellen O'Connor, Royal British Legion

RBL wishes to share findings from its 'Needs Analysis', a project designed to understand research gaps, and inform research commissioning, strategy, policy, and service delivery by drawing together the most comprehensive picture ever compiled, of the profile and needs of the armed forces community (AFC).

Census data was pivotal to improving our understanding of the AFC's demographics, needs, and service provision gaps. This was made possible by RBL's Count Them In campaign, which successfully introduced a question on previous military service into Census 2021 (England & Wales) and Census 2022 (Scotland).

This analysis has shaped RBL's research strategy by identifying priority areas for commissioned research to understand needs and how to meet them, including recovery services, finances, criminal

justice, and families' experiences. It has also informed RBL's policy priorities on issues including housing, employment, and healthcare. We are building an interactive dashboard using Census and other national datasets to present information about the AFC at national and local level, to inform and influence political decision-makers.

The analysis has enhanced service delivery by identifying areas of additional or unique need, allowing RBL to refine support services and expand provision in underserved locations. By cross-referencing Census insights with internal beneficiary data, we have pinpointed and filled gaps, such as the geographical distribution of RBL's Admiral Nurses (dementia specialists).

Session 6: Future of the census

Recent international census developments and their potential impact on the 2031 UK Census
Ian White, International Census Consultant

Every ten years the Conference of European Statisticians prepares a set of recommendations on the methodologies to be adopted by countries in the United Nationals Economic Commission for Europe (UNECE) Region for collecting and presenting information from the decennial population and housing census. The set of recommendations for the next (2030) round of international censuses - scheduled to be published later this year – include some significant changes in the direction in which countries may choose to collect their census data and the details of the data items to be collected and presented to users in outputs.

Among countries in the UNECE Region the move away from the traditional approach to census taking (that is, collecting the information directly from households using paper and/or electronic or online questionnaires – the approach adopted in the UK since 1801) to the use of the information already recorded on administrative registers and data sets (thereby obviating the need to contact the general public) is gaining pace. The majority of countries in the UNECE Region now expect to adopt this method of data collection, in varying degrees, in the 2030 round. There are some significant benefits to census takers (but also some challenges to users) in doing so.

Regardless of the method of data collection, changes to the recommended topic content of the census may have an impact on the way that data users can compare the information collected in the next round with the outputs from the 2021 and earlier censuses in the UK. The paper outlines some of the key changes in international recommendations and assesses the impact on the use of the resulting outputs in the next UK Census.

Reflecting on the future of the census
Oliver Duke-Williams and Nicola Shelton, University College London

The most recent censuses taken in the United Kingdom – in 2021 and 2022 – were also amongst the most unusual due to the circumstances of the pandemic. They may also be the last census of their kind: alternative data sources offer the potential for different ways of counting the population, bringing with them their own opportunities and threats.

This paper explores the future of population statistics in the UK based on the authors' experiences as the Census Service Director at the UK Data Service (ODW), and as director of the Centre for Longitudinal Study Information and User Support (NS). This includes the opportunities and threats posed by a move to make greater use of administrative data both for current and future data series including aggregate and longitudinal data, and longer term questions about how and why we preserve data for the future.