

Reflections from the UK on the future of population statistics

Oliver Duke-Williams¹ and Nicola Shelton; UCL, United Kingdom

This paper explores the future of population statistics – primarily in the UK – based on the authors' experience as the Census Service Director at the UK Data Service, and as a member of the census Methodological Assurance Review Panel, and as director of the Centre for Longitudinal Study Information and User Support (comments reflect personal perspectives). The topic of the presentation is thus census and other population data collection, and the data used include public statements and working papers from census agencies, as well as analysis of user behaviour. The findings are recommendations and observations for census data collection in the context of potential transformation to make more use of administrative data.

The 2021/22 UK censuses took place during the covid-19 pandemic; these circumstances were shared by many censuses around the world, but experienced uniquely within each country. These circumstances were however not wholly new: the 2001 UK census, for example, took place during a national outbreak of foot and mouth disease, requiring census enumerators in certain areas to adapt their work to meet necessary biosecurity requirements (ONS 2001). The potential problems faced were significant, and led for example to the postponement of the planned 2001 Census in the Republic of Ireland (CSO, 2002). Within the UK, the pandemic led to the postponement by one year of the (planned) 2021 census in Scotland, whilst the 2021 censuses in England and Wales and in Northern Ireland took place as planned, the differences reflecting differing local conditions and view on practicality of holding a census during the pandemic. Ignoring the implementation considerations, the 2021 UK census round was also interesting in a number of ways. It was the first set of set of censuses in the UK to be predominantly collected online (there had been an online option in 2011, with limited uptake) - again, this was a situation experienced in parallel in many countries, although some of those with quinquennial censuses had seen a growth in online completion in their 2016 rounds. More unique to the UK circumstances was that this was the first census conducted after the economic and social shock of Brexit, and was a population stock taking experienced amongst a flurry of anecdotal claims about people – especially EU citizens - entering or leaving the UK as a response to Brexit. The 2021 UK round saw the introduction of new questions – on sexual orientation, on gender identity, and on veteran status.

¹ Presenting / corresponding author

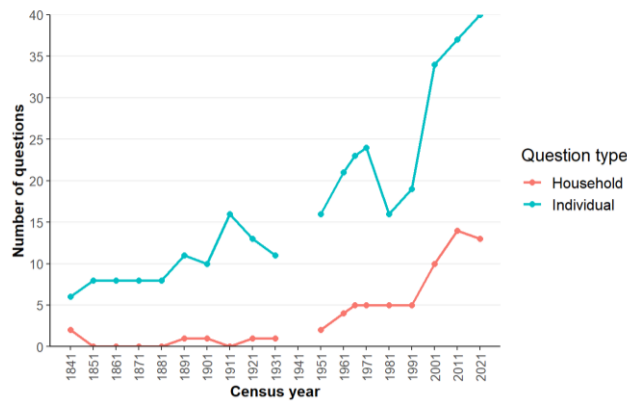


Figure 1 Numbers of census questions 1841-2021

The introduction of new questions is a common feature of UK censuses (Figure 1) and the paper will reflect briefly on the history of the census and on its growing complexity. However, the most pertinent aspect of the 2021 UK census round is the possibility that this will be the last such census in England and Wales, and perhaps throughout the UK. The suggestion that we are seeing the end of the census is nothing new, and has been signalled in previous censuses both in the popular media (for example, BBC 2011) and in academia (Coleman 2013, Martin 2006), although censuses have often somewhat stubbornly refused to die. The paper considers implications for current data assets.

Many countries are exploring the potential for the use of administrative data to replace in full or in part their traditional censuses. The approach to censuses is broadly similar in many countries – data are collected under the mandate of appropriate legislation, and processed and aggregated, and then published. Questionnaires are sometimes in paper format, sometimes online, often both, and included an integrated and internally consistent set of questions. Admin based population statistics are more unique to each country, made up of perhaps very different patchworked sources. This paper reflects on the opportunities and threats to population and migration data in England and Wales. On the one hand, the promise of admin data is significant. Vast amounts of data are already being collected by governments and their agencies, and these offer much more timely observations than those offered by censuses. Again, the 2021 census encapsulates the problems – whilst being ideally placed for those interested in the state of the nation in the middle of an extraordinary event, the timing was also such that it was roughly ten years out of date for those wanting detailed denominators for modelling exposure and populations at risk, etc. A more up-to-date set of denominators would be provided by annual (or more frequent) admin census outputs. However, admin sources are of course not perfect, and all tend to contain both overcounts and undercounts. A further concern is that not all of the themes covered in the census have obvious admin sources. In their consultation, ONS listed census variables classified by both coverage in admin data, and the

amount of research progress. A subset of variables were seen to have limited or partial coverage, and have had little or partial research progress. These include some frequently used variables. Table 1 is based on a submission to the consultation by the authors, and lists some of these 'variables of concern' (our description) together with the proportion of active research projects using the ONS Longitudinal Study (a linked census study) that make use of them. These vary in context: for some, such as caregiving, there are no obvious comprehensive sources, whereas for others such as education, there are good admin data (probably better in quality than census responses), but with some significant limitations; these will be explored more fully in the presentation.

Table 1 % of active LS projects (Q3, 2023) using selected census variables

Question / variable	% of active LS projects
Country of birth	68
Relationships in household	63
Educational level / qualifications	59
Occupation / social status	56
Self-rated health	44
Household tenure	37
LLTI or disability	32
National identity / ethnic group	27
Economic status / working activity	24
Welsh language use	20
Communal establishments	17
Caregiving	15
Transport available / distance travelled	12
Religion	12
Visitors staying over on census night	10
Passports held / main language / second address (each)	7

There are also challenges in data dissemination. The presentation will look at the experiences of the UK Data Service and other organisations in making available large quantities of census data for researchers, and challenges in supporting voluminous, complex, awkward aggregate data outputs from censuses from 1921 onwards, and of individual level records from almost a century earlier. How will such systems grow and adapt to provide access to admin sourced data, and how should they assist the researcher who is trying to do comparisons over time?

We can also consider the census in its much longer term context, with past censuses as a source for data for current historical demographers, and for genealogists and family historians, and current censuses as the source for their counterparts in the future. The sense of social contract – that we bequeath this data to the future – is rarely disputed in the UK, and in both the UK and the United

States activism is directed more at wider access to public records than to challenging the concept of future data release. This is not universal, and can be seen in response levels in countries such as Australia and Canada where there is – or has been – a requirement to actively opt-in to 'time capsule' retention and future release of data. This is not a given, and support has varied over time (Figure 2). The question of how we preserve born digital data for the future is an active one, and preservation over the timescale of a century is challenging: what will survive of us is unreadable data formats. A harder question on which the paper will reflect is one of what we should preserve of our mosaic of dry, ever changing, administrative data, and what might satisfy the analytical needs of future researchers, and offer the equivalent emotional heft of a census form to future family historians.

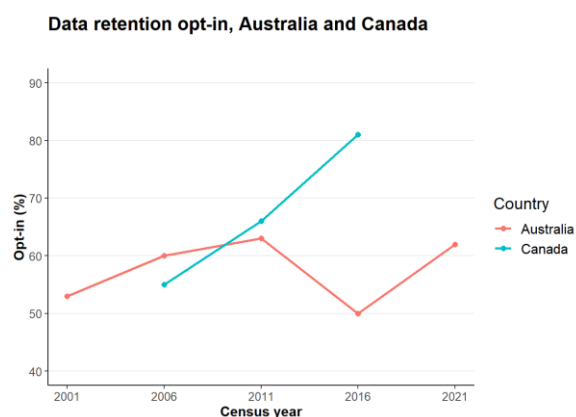


Figure 2 Census 'time capsule' opt-in rates, Australia and Canada

BBC (2011) Will the 2011 census be the last in the current format? <https://www.bbc.co.uk/news/uk-politics-13210080>

Central Statistics Office An Priomh Oifig-Staidrimh (2002) Census 2002 Preliminary Report <https://www.cso.ie/en/media/csoie/census/documents/2002preliminary.pdf>

Coleman, D. (2013). The Twilight of the Census. *Population and Development Review*, 38, 334–351. <https://www.jstor.org/stable/23655303>

Martin, D. (2006). Last of the censuses? The future of small area population data. *Transactions of the Institute of British Geographers*, 31(1), 6–18. <https://doi.org/10.1111/j.1475-5661.2006.00189.x>

Office for National Statistics (2001) 2001 Census Field Operations Report AG(01)04 <https://www.ons.gov.uk/file?uri=/census/2011census/consultationsusersandlocalpartners/censusadvisorygroups/operationalcag/ag0104tcm77190059.pdf>

Office for National Statistics (2023) The future of population and migration statistics in England and Wales https://consultations.ons.gov.uk/ons/futureofpopulationandmigrationstatistics/supporting_documents/Future%20of%20Population%20and%20Migration%20Statistics%20Consultation%20Document.pdf