International Journal of Population Data Science

Journal Website: www.ijpds.org

The Australian Census Longitudinal Dataset 2006-2016: A Multi-Wave Approach to Maintaining a Longitudinal Dataset

Kindermann, Bindi^{1*}, Nunan, Theresa¹, Hinde, Sarah¹, and Chipperfield, James¹

¹Australian Bureau of Statistics

Objective

The Australian Census Longitudinal Dataset (ACLD) is a nationally representative longitudinal dataset increasingly used to study Australian social and economic transitions. The ACLD will become an even more powerful resource when it is extended to include linked records from the 2016 Australian Census of Population and Housing.

The ACLD 2006-2011 dataset consists of a 5% representative sample of 2006 Census records and the linked 2011 Census records, and has been a well-received product in Australia as our first longitudinal Census dataset. In the process of extending the dataset to include linked records from the 2016 Census, we are aiming to improve the dataset by addressing issues such as linkage error and the under representation of specific population groups, while also taking advantage of recent changes in data linking methodology and policy that enable us to conduct better quality linking.

Approach

A multi-wave framework will be implemented for maintaining and improving the ACLD sample (as an alternative to the popular top-up sampling approach). Under the multi-wave framework, a sample is selected from each Census and is the beginning of a new wave of the ACLD, to be linked to each following Census. Allowing the sample selection method to be different for each wave enables us to apply learnings from previous waves to selecting a better sample for the next. For example, we could design the 2011 Wave sample to over sample certain population groups which were under represented in the 2006-2011 linked sample. We could also adjust the sample size, or change the method of sample selection in reaction to changing data item response rates.

Results

Initial analysis suggests the multi-wave approach will enable high quality analysis of transitions in hard to link and previously under represented population groups between 2011 and 2016. While these improvements cannot be retrospectively applied to the sample selected for the 2006 Wave, the linking strategy will also enable unbiased analysis of the Australian longitudinal population from 2006 to 2016.

Conclusion

The application of the multi-wave framework will enable the production of a high quality ACLD 2006-2016 product with minimal bias and improved sample representativeness, which can be used to explore changes in Australian society between the 2006, 2011 and 2016 Censuses. The flexibility of the multi-wave framework will allow continued improvement to the form and function of the ACLD in future waves.

*Corresponding Author:

Email Address: bindi.kindermann@abs.gov.au (B. Kindermann)