

# International Journal of Population Data Science

Journal Website: [www.ijpds.org](http://www.ijpds.org)



## New Zealand Longitudinal Census development

Kang, Inny<sup>1\*</sup>

<sup>1</sup>Statistics New Zealand

### Objectives

The aim for the New Zealand Longitudinal Census (NZLC) is to link data from seven recent censuses (1981-2013) and develop weight methodology to reduce the potential for link bias in the NZLC.

### Approach

The most workable way to create a longitudinal census dataset was to link adjacent pairs of censuses and then to link these census pairs together creating six linked pairs covering the seven censuses from 1981 to 2013. We refined the deterministic linking in the six pairs and then added probabilistic linking, where possible, for the remaining records in each of the theoretically linkable populations. The linking of the 2006 and 2013 Censuses was methodological challenge as the interval between censuses was 7 years instead of the normal 5 years between censuses. The challenge was heightened by having a question that asked for respondent address five years ago (ie in 2008).

In the linking of six censuses from 1981 to 2006, we had included indicators that identified completely linked and partially linked families. To link 2013-2006 census pairs, more variables were added in the probabilistic linking stage including geographic variables and new derived variables which included the combination of the role types within family group, family number, and sex variables. In developing weights for the linked pairs, all variables were changed into numeric values and then categorised. The data was then divided into male and female and logistic regression was used to predict probabilities of being linked.

### Results

The size of linked populations is relatively stable across similar time spans. We linked the majority of linked records in the deterministic stage which contributed links of for two-thirds of the theoretical populations. The subsequent stages included deter-

ministic linking and probabilistic linking which added approximately 5 percentage points to the link.

### Conclusion

The linking of census datasets from 1981 to 2013 to form a longitudinal data source allowed us to understand important aspects of population change across time which cannot be derived from simple cross. It provides a major new source of the analysis for insights into New Zealand demographic and social outcomes. Currently we are creating a set of generic weights for each individual record that may be applied in various research scenarios. Weighting of longitudinal data, as a part of NZLC, requires substantial in-depth statistical analysis of bias and coverage.

\*Corresponding Author:

Email Address: [innykang@gmail.com](mailto:innykang@gmail.com) (I. Kang)

