

## **Using address histories to improve the link rates of surveys in the Integrated Data Infrastructure (IDI).**

Nancy Wang

Statistical Analyst, Stats NZ

[www.stats.govt.nz](http://www.stats.govt.nz)

### **Abstract**

Stats NZ's Integrated Data Infrastructure (IDI) is a linked longitudinal database combining administrative and survey datasets. Initially, the IDI contained a small number of administrative datasets from key government agencies, which generally contained good quality identifying information such as names and date of birth. As a result, the methodologies developed to link these datasets together relied heavily on these variables, and yielded high link rates while maintaining good quality links. When survey datasets were later added, the link rates achieved were lower than that of administrative datasets, due to poor quality names and the underutilisation of geographic information. This indicated there were improvements to be made to the linking methodology used to link survey data in the IDI. Stats NZ underwent extensive consultation with the IDI research community on their requirements for the expansion of IDI (IDI2). A key finding from the consultation was that researchers wanted improved survey linkage in the IDI. This paper demonstrates how address history of individuals can be used to increase the link rate of surveys.