Abstract

Constructing household and family structures in linked administrative data in Australia

Clarke, C.[2], Buttsworth, S. [2], Gow, D. [2], Al-Janabi, S.[1], & Darby, P. [1].

[1] Life Course Data Initiative, People and Place Division, Australian Bureau of Statistics

[2] Data Linkage Methodology, Methodology and Data Science Division, Australian Bureau of Statistics

Increasingly, researchers are turning to linked administrative data for research purposes. Integrated data provides a longitudinal picture of a population, including pathways and touch points with government services. One current Australian data integration program is the pilot Life Course Data Initiative, which is led by the Australian Bureau of Statistics (ABS). This four-year program seeks to link administrative data in a careful and responsible way to enhance the longitudinal picture of child-wellbeing and outcomes. To make this new Life Course Dataset fit-for-use in addressing complex issues, such as child outcomes, it needs to be organised into household and family units. The aim of this presentation is to outline approaches for, and progress on, constructing household and family units with linked administrative data. Specifically, this presentation will: (1) provide an overview of the way in which households and family units have been conceptualised in various Australian data sources; (2) discuss different approaches for studying contemporary Australian families, especially from a child lens; and (3) outline approaches for, and progress on, organising administrative data into household and family units, using the Life Course Dataset as a case study.

Extended Abstract

Background

Over the past few years, many reports on disadvantage in Australia have been released by government, philanthropy and academics. These reports have provided informative statistics on the extent and depth of disadvantage in Australia, stressing the importance of establishing holistic policies, programs, and practices that help reduce disadvantage. Some of these reports indicated that the risk factors of disadvantage commence in childhood and acknowledged the importance of targeting disadvantage in the early years. Some of these reports also acknowledged the importance of integrated data and data analytics in helping provide an evidence-base for those programs, policies, and practices. The Life Course Data Initiative serves as one such current data integration program.

The Life Course Data Initiative

The 2023-24 Commonwealth (Australia) Budget delivered a $199.8 million package to target entrenched community disadvantage with a focus on intergenerational disadvantage and improving child and family wellbeing. The Treasury and Department of Social Services are leading the implementation of the package, with the support of the Australian Bureau of Statistics (ABS). Further information on the Targeting Entrenched Disadvantage Package is available. As part of the Budget commitment, the ABS received $16.4 million to deliver the Life Course Data Initiative, commencing July 2023. The Life Course Data Initiative is a four-year program that will create a dataset and associated statistics to inform responses to address child disadvantage, particularly that which is entrenched.

The Life Course Dataset

The Life Course Dataset will include new administrative data safely linked to the Person Level Integrated Data Asset (PLIDA), which is a secure data asset that includes information on health, education, government payments, income and taxation, employment, and population demographics (including the Census) over time. Additional data linked to PLIDA will include national child care data and birth registries data, Australian Capital Territory Government health and education data, and South Australia Government Health, Education, Housing, and Homelessness data from the Better Evidence Better Outcomes Linked Data Asset. This Life Course Dataset will, therefore, improve data coverage for Australian children, particularly those born from 2006-07, providing from the antenatal period, through to birth, and up to 14 years of age cutting across multiple domains

important to a child's development and outcomes. The Dataset will also include information on the families and households, and spatial communities, of children.

Household and family units

Household and family units in administrative data can be used to understand the experiences of a child and their support networks. This can aid understanding of issues relating to income support, social housing, housing stress, child disadvantage, household crowding, and household income. Hence, to make the Life Course Dataset fit-for-use in providing insights relating to child disadvantage, the purpose of the current work is to outline and test approaches for organising the Life Course Dataset into quality household and family units.

Theoretical framework

The way in which households and families are operationalised is critical for determining who counts. For the purposes of this work, households will be defined as persons, at least one of whom is at least 15 years of age, who reside in the same private dwelling (i.e., share an address). A household can contain multiple families and families can exist across households.

The definition of families will be broader to account for the experiences of contemporary Australian families, which may be geographically and temporarily dynamic with flexible boundaries for who counts. Families will be defined by identifying the relationships children have with people in the Life Course Dataset. This may include parents, guardians, siblings, or grandparents who are providing the child with support. These persons may or may not be residing with the child in a household.

Methodological approach

The construction of households will be based on a process of grouping persons with the same address (using the Australian Address Register ID, or ARID) in the Life Course Dataset at a certain point in time (annual snapshot). In instances where the ARID may be missing, inconsistent or may only be able to be coded to a broader geographic street address, we will explore the opportunity for inferring or imputing the address. The derived household units will be validated by using a variety of approaches, including comparing to known benchmarks and comparing to the Census.

The construction of families will be based on a process of identifying relationships to children in the Life Course Dataset using episodic structures, indicating best known start and end dates. The children in scope will be those born in 2006-07. Where, for example, relationship information may be missing for children, we will

explore the opportunity for inferring or imputing the relationships or estimating them using modelling approaches based on the Census.

Anticipated findings

Through this work, we anticipate identifying strategies for constructing household and family units that could be implemented using information contained within the Life Course Dataset. We recognise that these approaches may not adequately account for all types of family dynamics, such as those of Aboriginal and Torres Strait Islander peoples, or children who are in shared care arrangements. We will, therefore, also discuss the ways in which these units can be used by researchers and known gaps or data quality issues that can be ameliorated in future work.