

Preparing for low response events in the Australian 2026 Census

Ian Rayson, Australian Bureau of Statistics (ABS)

Summary

The Census is one of Australia's largest peacetime operations. Each Census has a risk of low response – Stats NZ 2023 Census felt the impact of Cyclone Gabrielle, Stats Canada 2016 Census had disruptions from a bushfire in the town of Fort McMurray, and Stats NZ 2018 Census experienced unexpectedly lower response rates across the country. While Australia has remained largely immune, it is important to be prepared to fill gaps in population data when low response events happen.

The Australian Bureau of Statistics (ABS) existing method to fill in gaps in Census data is good when Census participation is high, however is less robust in low response scenarios. The ABS is exploring a new method in the 2026 Census. This method would use PLIDA data directly to fill in gaps in data where households did not return a Census form.

There are several benefits from this ADS approach:

- These records will more accurately reflect people missed in the Census, with data sourced directly from administrative records.
- It has the potential to fill in more data across a greater range of characteristics about people missed in the Census, including potentially reducing the undercount of Aboriginal and Torres Strait Islander persons.
- It is a stepping stone for potential greater use of administrative data in future Censuses.

This paper will present background information on the Census and existing methods used to fill gaps. It will then describe the new proposed method for 2026, and demonstrate our preparedness for filling data gaps in areas that experienced poor response rates in the 2021 Census.

Theoretical focus

Every 5 years, the Australian Census, run by the ABS, aims to count every person and household in Australia. This is done primarily by contacting households via letters or visits, then households complete the form and submit it online or send it back in the mail. The ABS invests considerable effort to maximise the response rate of the Census, including targeted follow up of households that have not returned a form. Despite this, and although Australia has one of the highest participation rates in Censuses worldwide, we do not receive a form from all households – in the 2021 Census 96.5% of households responded to the Census.

To ensure the Census count accurately reflects the population in Australia, ABS fills in gaps in data left by people from these non-responding households. In 2021 this was done by an imputation method known as hot-decking. This involves randomly selecting a donor record and copying the donor's responses into the household requiring imputation. The donor record is designed to have

similar characteristics and be located geographically as close as possible to the location of the household to be imputed. In 2021, administrative data was used in hot-deck imputation for the first time to improve the selection of donors.

Although the hot-decking imputation method works well when Census participation is high, it is less robust in low response scenarios. In the next Census in 2026, the ABS proposes to further improve the imputation of non-responding households by using administrative data directly, where it is available. This direct administrative data use for imputation is known as Admin Data Sourcing (ADS). Not all Census non-responding households have persons recorded in administrative data however, in which case the existing hot-decking imputation method would be used.

In most low response scenarios, while the hot-decking method does not adequately fill gaps, ADS is expected to adequately fill the larger gap left in the Census in a low response scenario. There are situations however where ADS will not adequately fill gaps in low response scenarios, for example in regions with poorer population coverage of administrative data, or where people's location cannot be adequately recorded at an address level. In these cases, bespoke methods may be required to adequately fill remaining gaps. These bespoke methods vary depending on the nature of the low response event, and the administrative data available in the area experiencing low response. Examples of bespoke methods are described in the Research Methods section of this paper.

The existing hot-decking imputation method only imputes basic characteristics about the people in the non-responding household. It imputes Age, Sex, Marital Status, and Place of Usual Residence. It does not impute the full range of information collected in the 65 questions on the Census form, and of note, Indigenous Status is not imputed. This approach leaves a large untreated undercount of Aboriginal and Torres Strait Islander persons in the Census, which in 2021 was 17.4% nationally (compared with a net undercount of 0.7% for all persons after imputation). In low response scenarios this undercount is even larger, and is even more pronounced in areas with a predominantly Indigenous population.

One of the key benefits of ADS is that it can impute a wider range of characteristics than hot-deck imputation, including Indigenous Status. Subject to Indigenous stakeholder consultation, if Indigenous Status is included in the information used by the ADS method, it is expected to substantially reduce the Indigenous undercount in the 2026 Census, and even more so in low response scenarios.

Data

To fill Census data gaps left by households that did not respond to the Census, the ADS method uses data similar to the [Administrative data snapshot of population](#). This dataset, which was published as an experimental statistical dataset in August 2023, is an extract of government integrated administrative data, called [PLIDA](#), which is scoped to an Australian resident population around Census time. Nationally, the population count in the Administrative data snapshot of population is within 0.2% of Australia's official population count.

The PLIDA data includes usual residence address information from Tax, Medicare, and Social Security data, and thus can be used to fill gaps in Census data due where no data exists for that household.

PLIDA includes a wide range of characteristics that are not used in the hot-decking method, notably Indigenous Status and Income. These present an opportunity to improve the coverage of these data items in Census data.

Research methods

For households that did not return a Census form, the ADS method uses records from the Administrative data snapshot of population, as described in the Data section of this paper, where it's available.

A person's address in the admin data snapshot may not be the same as where the person lived on Census night. In this situation if we used the ADS data directly, we would double count these people, as they are included both on a Census form and in the ADS. To mitigate against this we have a deduplication step, where we remove people from the admin data snapshot who have already responded to the Census.

Not all addresses in the Census have corresponding data in PLIDA. The existing hot-decking imputation method will be used where an address has no persons in administrative data recorded there.

Extensions of the ADS method are being explored in case of situations where areas have severely low response, and further Census field work and ADS do not adequately fill the gaps. These include:

- Manual reviews of address text to find further street addresses
- Using relationship information to group people together into households, even if they don't have a codable address
- Filling gaps in Census data due to people missed in responding households
- Filling gaps at an area level, where PLIDA information is available about people living in that area but are unable to be located within an address or household.

The ADS and extension methods will be researched using 2021 Census data as a base. The ADS method will be applied firstly to the full Census responding dataset – a normal response scenario. Results will be assessed at a National, State, and area level; with emphasis on areas that had a lower response rate. We will then conduct a low response simulation study, where further households are removed from the Census dataset and the ADS applied. The Census count and demographic features using ADS will be compared to those that would have been produced if hot-decking imputation had solely been used.

The research will focus on case studies from the 2021 Census where local areas experienced a much lower response rate. In these areas we will investigate the extent to which the ADS method and relevant extensions can fill the gaps.

Expected findings

The ADS is expected to improve upon the existing imputation method. Using a case study on 2021 Census data, in a normal response scenario ADS would be used for about 45% of non-responding households. In these households, the ADS records will more accurately reflect people missed in the Census, with data sourced directly from administrative records.

We expect that Admin Data Sourcing will adequately remediate Census data in a national low-response situation, like what was experienced in the 2018 NZ Census. We will present findings at a National and area level where an across-the-board low response scenario is simulated.

ADS has the potential to fill in more data across a greater range of characteristics about people missed in the Census, including potentially reducing the undercount of Aboriginal and Torres Strait Islander persons. A decision to proceed with this will be subject to extensive Aboriginal and Torres Strait Islander stakeholder and community consultation. Based on 2021 data the ADS would assist in reducing the Census undercount from 17% to 11%.

More crucially, the ADS embeds the direct use of admin data within standard Census data processing systems, meaning that the ABS is well prepared for the greater usage of administrative data will likely be required to remediate Census data in the event of a low response. Looking beyond the 2026 Census, ADS is a stepping stone for potential greater use of administrative data in future Censuses.

We will present several case studies of local areas that experienced a low response in the 2021 Census. This includes some areas with a relatively large Indigenous population, where the count of Indigenous persons is expected to substantially improve. We expect that the ADS will fill gaps in Census data to some extent, but in some isolated cases will still leave substantial gaps. We expect that the additional low response contingency methods will fill remaining gaps adequately, ensuring Census provides accurate local level data for these local areas. We expect that the specific method to fill gaps in a particular local area will be bespoke. It will depend on the nature of the low response event at Census time, and on the quality of administrative data in the area.