



Unlocking the full potential of nationally-important datasets

Christine Williams, Education and
Data Integration Branch



Outline

- Data Integration - unlocking the full potential of datasets
- ABS as an accredited integrating authority
- ABS data integration projects (examples)
 - Census Data Enhancements – Unlocking the full potential of the Census
 - Creating new longitudinal datasets



Unlocking the potential of data

- Maximising the use of existing datasets
- Opportunities to use new data sources (including Big Data)
- Innovative and strategic solutions to complex problems



Data integration at the ABS

- Vast data holdings
 - Economy
 - Environment
 - Social
- Experience handling sensitive information in an effective and safe way
- Commonwealth accredited Integrating Authority



ABS model for public trust

- Legislative framework
- Privacy impact assessment
- Strong security (physical and electronic)
- Established data management practices
- Audit mechanisms (electronic and manual)
- Separate server environments
- Disclosure control
- Organisational culture of confidentiality
- Analytical capability



Confidentiality

- Physical and electronic security of data
 - from collection through to archiving
- Culture of data management
 - ‘Need to know’ access to data
 - Protection of *Census and Statistics Act 1905*
- Confidentialised dissemination of data
- Transparency
- Continued improvement and investment



ABS data integration projects

- Census Data Enhancement projects:
 - Death Registrations to Census Linkage Project
 - Education and Census Quality Study
 - Migrants Census Data Enhancement Project
 - Statistical Longitudinal Census Dataset
- Creating new longitudinal datasets:
 - Statistical Longitudinal Census Dataset
 - Business Longitudinal Database
 - Longitudinal Labour Force
- PIT data: Migrant Personal Income Tax Project



Census Data Enhancement

- Integrating unit record data from the Census with other datasets
- Maximises the utility and enhances the value of the Census
- Key benefits realised:
 - Improved accuracy of Aboriginal and Torres Strait Islander life expectancy
 - Creation of a longitudinal Census dataset (SLCD)
 - Improved information on migrant outcomes



Death Registrations to Census Linkage Project

- Improve the measurement of Aboriginal and Torres Strait Islander life expectancy
- Links 12 months of death registrations to the 2011 Census
- Around 80% of Aboriginal and Torres Strait Islander death records linked to their Census record
- New life expectancy estimates to be released on 15 November 2013



Census Education Quality Study

- Linked school enrolment records (from QLD, SA, Tas, and NT) to 2011 Census data
- Feasibility study analysing four linkage standards
- Over 85% of enrolment records linked for QLD, SA and Tas
- 76% of NT enrolments linked



Key Elements of the Migrants CDE

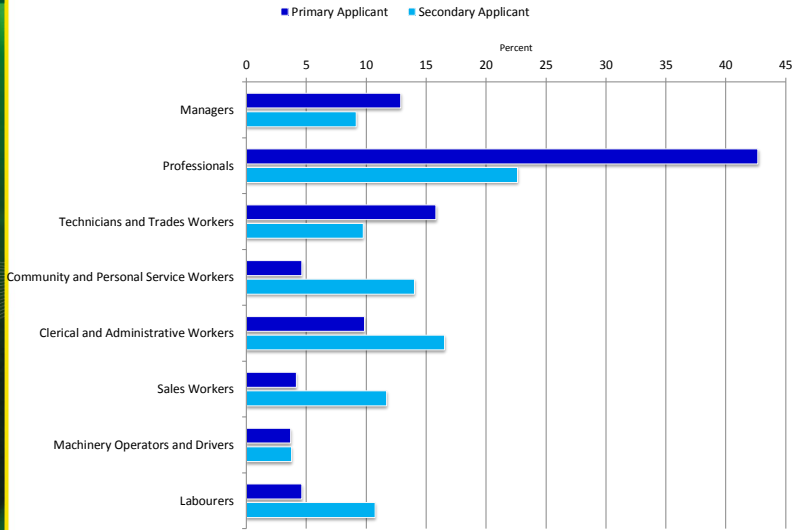
- Linking the Census file to the Settlement Database (SDB)
- Quality study published August 2013
- Statistical study published September 2013
- Future plans to link SDB to SLCD for longitudinal information
- 76% of Settlement Database records linked to Census records



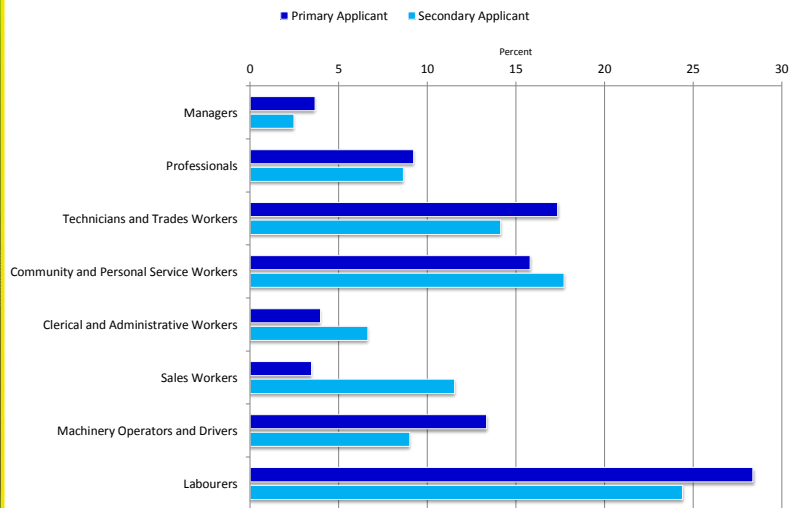
Statistical Results of Migrants CDE project. . .



Skilled migrants, occupation by applicant status (age 15+)

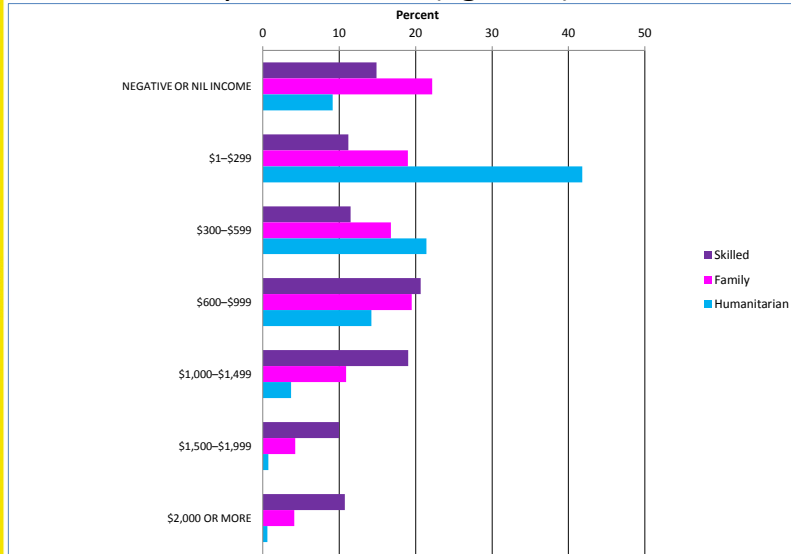


Humanitarian migrants, occupation by applicant status (age 15+)





Personal weekly income of permanent migrants by Visa stream (age 15+)



Statistical Longitudinal Census Dataset (SLCD)

- Understand how social and economic conditions change over time at the individual level
- Analyse changes to populations of interest over time
- Opportunity to expand the dataset through data integration
- 2006-2011 SLCD released December 2013



SLCD

Possible topics for analysis:

- Youth and children outcomes
- Lifecycle pathways and decisions
- Migrant experiences
- Pathways to retirement
- Family formation and the impact of divorce/separation



Personal Income Tax data in the ABS

- ABS is currently the only agency to have a partnership with ATO to access Personal Income Tax data for statistical and research purposes
- The potential of PIT data:
 - longitudinal analysis
 - labour market and macroeconomic statistics
 - regional statistics
 - other integrated data



Migrant Personal Income Tax Project

- Feasibility study integrating the Migrants Settlements Database (DIAC) with Personal Income Tax Data (ATO)
- Will produce a case study for probabilistic linkage between administrative data
- If successful, future statistical studies will be proposed



Further information

For more information about ABS data integration services and activities visit:

www.abs.gov.au/dataintegration

Or email:

data.integration@abs.gov.au