

Slide 1



Hi everyone,

In this first session I will be describing very briefly a number of data sources available at DSS. The notes from these presentations will be available afterwards on the community grants hub website for your reference, so please don't feel that you need to remember everything!

Slide 2

How might this be relevant to your grant application for the TTL Fund?

The assessment criteria outlined in the Try Test and Learn Fund grant opportunity guidelines encourages you to provide 'evidence'

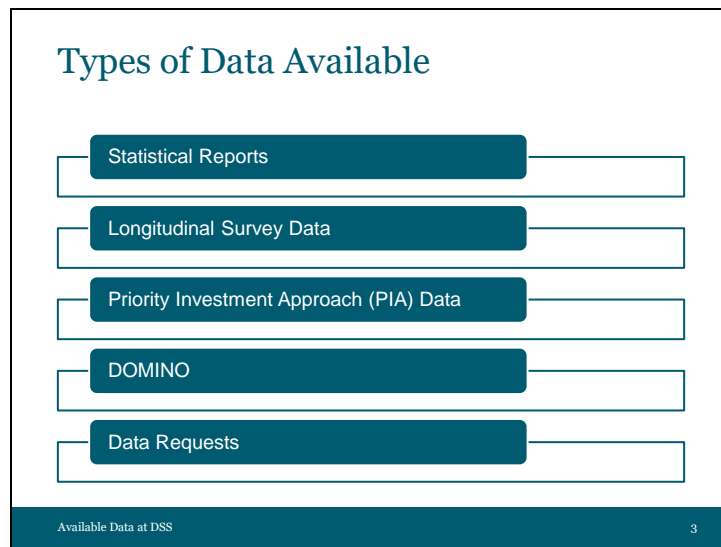
- provide evidence that the people targeted by your project are at risk of long-term welfare dependence (e.g. evidence may include Priority Investment Approach data, research, government reports, empirical evidence, etc.);
- provide evidence of the need for your project among those it would support (e.g. evidence may include Priority Investment Approach data, research, government reports, empirical evidence, etc.)

An important place to start this session is to consider how this information might be relevant to you and in particular for your grant application for the TTL Fund.

The assessment criteria outlined in the Try Test and Learn Fund guidelines encourages you to provide 'evidence'.

- Both of these examples refer to the Priority Investment Approach data as well as research and government reports. Your evidence is likely to include your own information and data from your experience. However, your information can be given greater context and supported by the data that is now available to you from a range of DSS data access points.

Slide 3



I will cover these five avenues of information that are available:

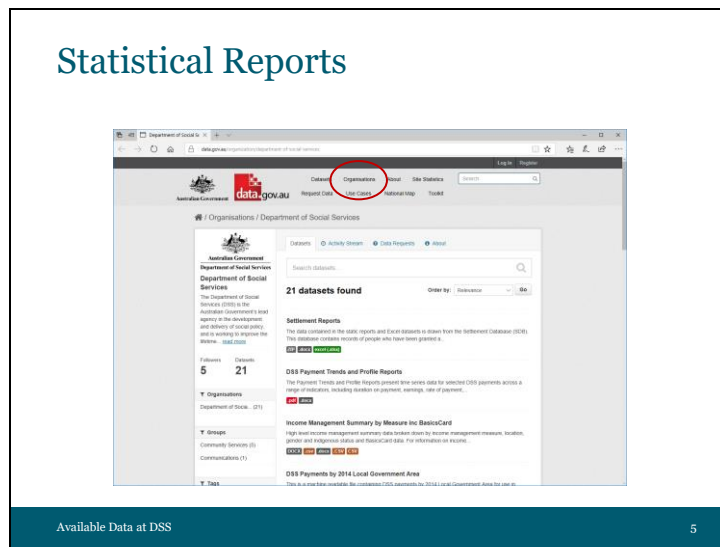
- Quarterly Statistical Reports
- Longitudinal Survey Data
- Priority Investment Approach Data, or PIA
- DOMINO and
- Data Requests

Slide 4



The quarterly statistical reports are the first point you might go to for general information.

Slide 5



Data.Gov holds a large amount of publicly released data across all government agencies. To find the data we have available on Data.Gov, please search

- Organisations for 'Department of Social Services'


When you have come to the DSS page you will be able to scroll down and see the 21 datasets held there. One of these is the Payment Demographic Data. This dataset presents information about people on a range of benefits.

Statistical Reports

Payment type	September 2017					September 2017				
	16-20	21-24	25-34	35-44	45 and over	16-20	21-24	25-34	35-44	45 and over
16-20	1,546	287	70	549	144	40	71	64	2	4
21-24	286	496	84	38	20	3,824	186	84	21	2,814
25-34	0	20	34	662	777	414	834	475	2,048	14
35-44	0	0	40	40	12	40	0	0	0	0
45 and over	0	0	0	0	0	0	0	0	0	0
Total	1,832	783	144	1,993	1,667	8,198	2,476	1,529	2,073	2,822

The reports allow you to view the number of people receiving a benefit in that particular quarter. You can look at a number of factors such as age, gender, language, country of birth and you can see this by location.

Slide 7



Longitudinal Surveys
**National Centre for
Longitudinal Data**
www.dss.gov.au/nclld
Available Data at DSS

7

Longitudinal data is quite different to the administrative data normally made available by a government department. With longitudinal data, information is collected from the same individuals repeatedly over time. This means that we are able to see how actions and events can affect outcomes over time.

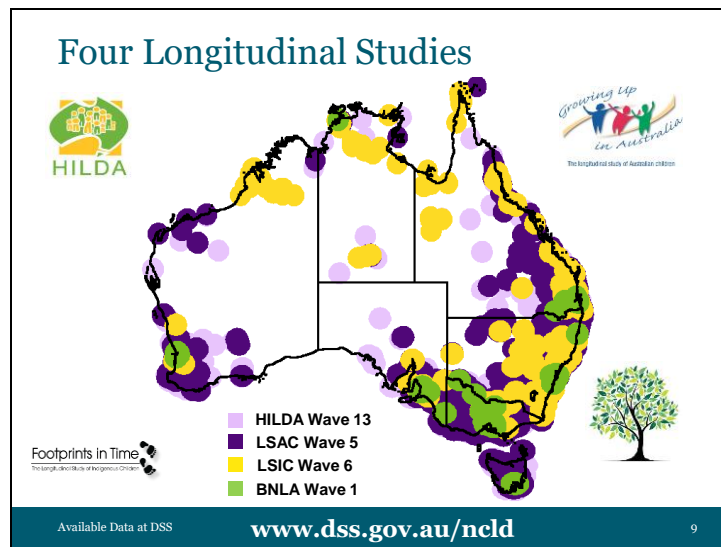
Longitudinal Surveys: National Centre for Longitudinal Data

- The Household, Income and Labor Dynamics in Australia Survey (HILDA)
- The Longitudinal Study of Australian Children (LSAC)
- The Longitudinal Study of Indigenous Children (LSIC)
- Building a New Life in Australia (BNLA)

Available Data at DSS 8

Through the National Centre for Longitudinal Data, or NCLD, DSS runs four distinct longitudinal research studies each with a different sample population and different purpose. These interviews are run every year or every second year and often take approximately two hours. These participants are voluntarily providing their time and information each time. The interviews ask about a range of topics that include quite personal experiences and opinions.

Slide 9



To give you an idea of the geographical areas these studies cover, let's look at the distribution of each.

- You can find information on each of these studies on the NCLD website. You will also find the reports that publish data from the studies, fact sheets and a link to where you can access the data. I have a copy of the fact sheets here for you to have a look at.

The Household, Income and Labour Dynamics in Australia (HILDA) Survey

WHO?

- More than 9,000 households
- More than 17,000 individuals



WHERE?

- Representative of urban and regional areas across Australia

HOW OFTEN AND FOR HOW LONG?

- Surveys are run every year and started in 2001

HILDA, the study of household, income and labour dynamics, follows more than 17,000 individuals. This study started in 2001 and interviews the participants each year. The sample is representative of urban and regional areas with some participants who live remote.

The Household, Income and Labour Dynamics in Australia (HILDA) Survey

WHAT IS IN IT?

- Household, budget and income information as well as family information such as caring, health and disability, and personal information such as education, wellbeing and relationships.

WHERE CAN I GET IT?

- www.dss.gov.au/nclid

HILDA includes household, budget and income information, family information such as caring, health and disability, and personal information such as education, wellbeing and relationships

Growing Up in Australia: The Longitudinal Study of Australian Children (LSAC)



WHO?

- 10,000 children originally recruited

WHERE?

- Representative across Australia by population density

HOW OFTEN AND FOR HOW LONG?

- Surveys are run every second year and started in 2004

LSAC, the study of Australian Children, originally recruited 10,000 children in 2004 including 5,000 babies and 5,000 children from Kindergarten. The babies are now 14-15 years old and the kindy kids are now 18-19 years old. This study started in 2004 and interviews the participants every second year.

Growing Up in Australia: The Longitudinal Study of Australian Children (LSAC)

WHAT IS IN IT?

- Family relationships, home environment, health and wellbeing, education, work and care responsibilities, risk behaviour and intergenerational information
- Some linked data


WHERE CAN I GET IT?

- www.dss.gov.au/nclcd

LSAC follows the children but also interviews the primary parent. This means that there is information from both perspectives in terms of family life and factors that might influence the child's development. Some types of information are able to be compared between the parents and children such as education and risk taking behaviour.

There are a number of linkages that have been made such as NAPLAN results, Centrelink data and MBS/PBS data.

Footprints In Time: The Longitudinal Study of Indigenous Children (LSIC)



WHO?

- 1,700 indigenous children

WHERE?

- Surveys conducted in remote, rural and urban locations across Australia

HOW OFTEN AND FOR HOW LONG?

- Surveys are conducted every year and started in 2008

Available Data at DSS 14

LSIC, the study of indigenous children, is following 1,700 indigenous children and also started with both babies and children in kindergarten. These interviews are conducted every year and the children are now 10-11 and 13-14 years old. This study started in 2008 and interviews the participants each year. Remote areas are well represented in this study as well as rural and urban locations.

*Footprints In Time: The Longitudinal Study
of Indigenous Children (LSIC)*

WHAT IS IN IT?


- Family relationships, housing, education, cognitive development, health, employment, connection with indigenous culture, and experience of racism

WHERE CAN I GET IT?

- www.dss.gov.au/nclid

LSIC includes information on family relationships, housing, education, cognitive development, health, employment, connection with indigenous culture (including language and country), and experience of racism.

Building a New Life in Australia: The Longitudinal Study of Humanitarian Migrants (BNLA)



WHO?

- 1,500 migrating units

WHERE?

- Large and small communities around Australia

HOW LONG?

- Surveys are conducted every year and started in 2013

Available Data at DSS 16

BNLA, the study of humanitarian migrants, follows 1,500 migrating units (the grantee of a humanitarian visa plus others – typically family members). This includes approximately 2,400 individuals who were granted a permanent humanitarian visa in the 3-6 months before the study started. The sample ranges in ages from 15-83 yrs old. Participants are living in large and small communities around Australia, across most states and territories, with most living in Melbourne and Sydney. Almost all participants arrived in Australia in 2013 to start a new life so this study follows their experience from the early days of their time here.

*Building a New Life in Australia: The
Longitudinal Study of Humanitarian
Migrants (BNLA)*

WHAT IS IN IT?

- Housing and resources, language, education, employment, health, access to services, community support and their life experiences before coming to Australia

WHERE CAN I GET IT?

- www.dss.gov.au/nclcd

BNLA includes information on housing and resources, language, education, employment, health, access to services, community support and their life experiences before coming to Australia.



The original Priority Investment Approach data has been used internally to inform the development of the Try, Test and Learn fund. This is the dataset that has been used to identify the priority groups outlined in both Tranche 1 and Tranche 2. A subset of this data has been made available for researchers and a broader audience.

The three access points to PIA data provide tiered access to enable greater accessibility to a wider audience while maintaining privacy

	PIA in SURE	TableBuilder	Synthetic Data
Accessibility	Applications are considered and approval based on user expertise and need for the data in context of a specific project	Unrestricted access	Unrestricted access
User Access	User pays an access fee	Free registration	Free registration
User Experience	Once approved, access is within a secure enclave and users will see unit record data. All analyses are run remotely within the secure enclave. Output must be approved before being taken out of SURE	Once users log in, they are able to construct tables of aggregate data freely, without constraint. These tables can then be downloaded as .csv files without review. Users do not see the unit level record data at any stage	Users can download the complete dataset of unit level record data to use on their own system
Sources	AIHW secure environment	ABS Tablebuilder site	Dataverse
Number of Variables	60	38	31
Format of Date of Birth	Rounded to the month	Rounded to the year	Rounded to the decade
Geographic Information	Country (Birth, Residence and Citizenship), State, Postcode	Country by Region (Birth and Residence), State, Postcode, Address at SA3 level	Country (Birth and Residence), Postcode

Access to Priority Investment Approach (PIA) Data 19

Part of the PIA policy includes providing access to the data to help improve the lives of Australians who are receiving welfare. A tiered access model was developed so that detailed information is available to experienced researchers in the secure environment (or SURE), while less detailed information is available to a broader audience. This approach maintains the privacy and security of the data while making the data as useful as possible.

As TableBuilder and the Synthetic Data are publicly accessible, a number of steps have been taken to ensure privacy is maintained.

I would like to point out

- the relative accessibility of the Tablebuilder and Synthetic Data (i.e., free registration and unrestricted access) and
- also some of the key differences in how the data has been made confidential for privacy purposes. For example, you can see there are less variables in TableBuilder and Synthetic Data compared to the PIA data that is available in the Secure Enclave of SURE and that
- the date and geographic information is less specific. For example, country of birth is only shown by region in TableBuilder whereas the specific country is shown in SURE, and DOB is by month and year in SURE where it is year in TableBuilder and decade in the Synthetic Data

Priority Investment Approach (PIA) Data

- The PIA Research dataset is a subset that includes 14 years of information about payment recipients over 56 quarterly snapshots from July 2001 to June 2015.
- The data contains information about those people who have received one of 21 benefits including Aged Pension, Disability Support Pension, Newstart Allowance, Youth Allowance and Carer Payment.
- The de-identified information includes demographic and geographical information, as well as information relating to accommodation, primary medical conditions, education and income.
- Demographic variables of the partner are also available for some of the recipients.

The PIA Research dataset includes information about payment recipients over 56 quarterly snapshots from July 2001 to June 2015.

- The data contains information about those people who have received one of 21 benefits including Disability Support Pension, Newstart Allowance and Youth Allowance.
- The de-identified information includes demographic and geographical information, as well as information relating to accommodation, primary medical conditions, education and income.
- Demographic information about the partner is also available for some of the recipients.

PIA Data in SURE (and the Synthetic Data)

AUDIENCE

- Researchers with advanced analytics skills

FUNCTION

- Unit-record level data for research purposes

LOCATION

- www.aihw.gov.au/our-services/data-on-request/dss-data

Available Data at DSS 21

The PIA data in the secure enclave is available to researchers with advanced analytics skills on an as-needs basis. Each application needs to be made identifying the specific variables required to address a specific research question.

All analyses are conducted within the secure environment and output must be approved prior to being taken out.

There is a fee payable for each application to recover the costs for providing the service.

The advantage of the design of this release is the data is already loaded and available. This means applicants experience a stream-lined process with a much-reduced wait-time.

The Synthetic Dataset is available for researchers to test their models before accessing the PIA data in SURE. As it is not particularly relevant to this forum I won't go into it here but if you would like to know more about it please come and see me in the break.

PIA Data in TableBuilder

AUDIENCE

- Everybody
 - Teachers and students
 - Academics
 - Non Government Organisations

FUNCTION

- Create and download both tables and graphs
- Population estimates

LOCATION

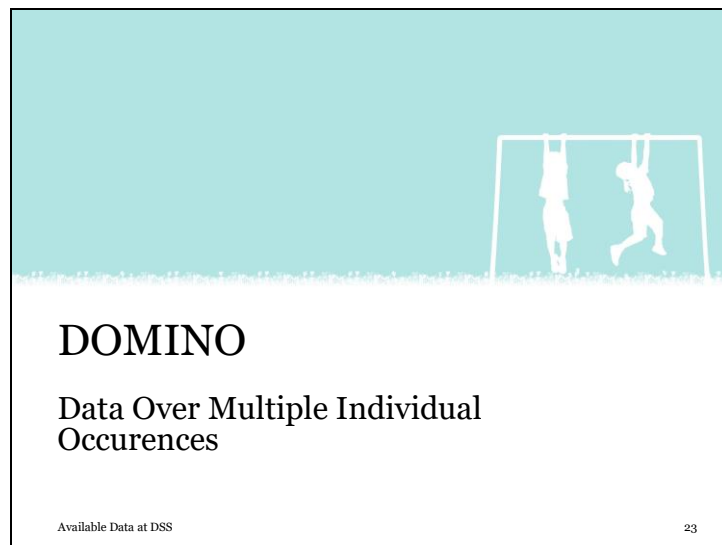
- Australian Bureau of Statistics

Available Data at DSS 22

With the PIA data in TableBuilder, users are expected to include teachers and students in an educational context, as well as academics and NGO's looking for data to use in their grant applications.

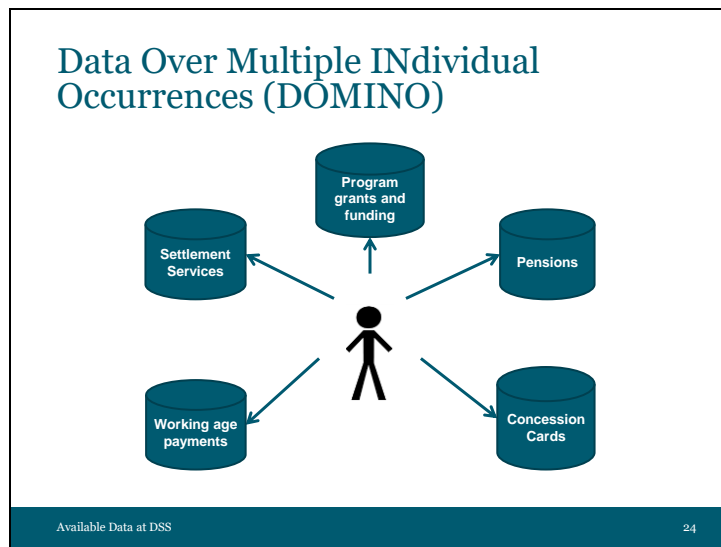
TableBuilder uses a 5% sample of the research PIA dataset available in SURE. This is large enough to provide a representative subsample of the whole population while maintaining individual privacy. One implication of this is that if you are reporting population estimates in your grant application, you will need to multiply the figures in the table by 20.

The next session demonstrates what you can do with TableBuilder and how to get there.



Typically administrative data has been collected as quarterly snapshots as is the case with the PIA data and the Statistical Reports.

DOMINO is a new dataset that has taken a very different approach to looking at administrative data.



Historically, DSS have had a wealth of data, over multiple programs. Through years of programme changes, and new system creation, this data has become 'siloes'. While this allows for an individual programme to have very detailed data, it is very hard to track customers across programs and services or through major life events and changes. Being able to do this would allow services to better assist in creating 'lifetime wellbeing'.

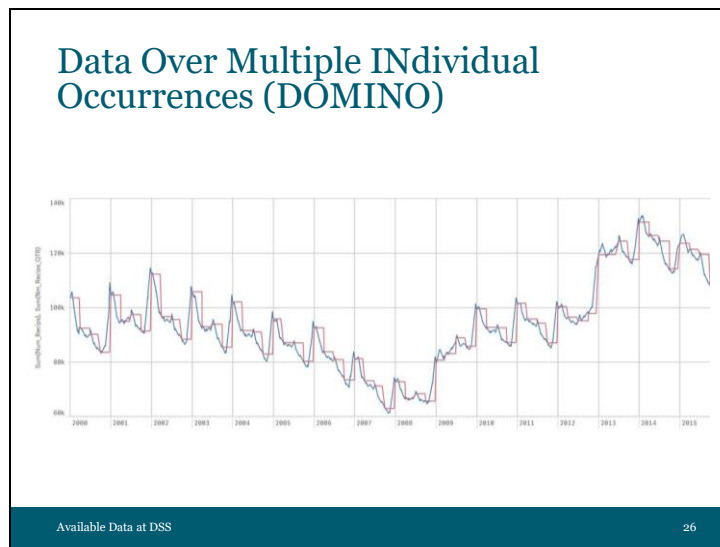
Data Over Multiple INdividual Occurrences (DÓMINO)

- Integrated Research Data Mart of DSS Administrative Analytical Data Assets.
- Event-based structure that tracks changes in individuals circumstances over time and interactions with both Social Security Payments and DSS managed programs.
- Enables tracking of individual customers as they interact with DSS longitudinally throughout their life course.

Available Data at DSS 25

DOMINO is not a single dataset, but instead, a DataMart – each table similar to a domino ‘piece’, made to be linked to other pieces.

- The event-based structure tracks changes in individuals circumstances over time and interactions with Social Security Payments and DSS managed programs.
- This enables tracking of individual customers as they interact with DSS longitudinally throughout their life course. All individual records are de-identified but the longitudinal nature of the data is retained.



**** QUERY: Is this slide acceptable to use for this presentation, regarding PIC/data security, etc? If not, delete this slide altogether.**

The new approach of looking at all events, rather than counting customers at a single point in time provides a new way of viewing data.

The chart above shows Youth Allowance (other) mapped on two distinct lines.

- The RED line is historically how we view data, based on quarterly snapshots. While it shows peaks and troughs in payment, it doesn't have the dramatic picture of how long each peak lasts.
- The BLUE line is using DOMINO and daily events. This line roughly follows the quarterly snapshot, but provides a much more detailed story of how the population changes over time.

DOMINO Data in SURE

AUDIENCE

- Researchers with advanced analytics skills

FUNCTION

- Unit-record level data for research purposes
- Designed for linkage studies

LOCATION

- www.aihw.gov.au/our-services/data-on-request/dss-data

Available Data at DSS 27

DOMINO is also in the secure enclave with the PIA data. Due to the complexity of the data, it is only available to researchers with advanced analytics skills.

DOMINO has been designed to enable data linkage projects. If this is of interest to you, please come and see me in the break or email in after the session.

Slide 28



We also have a service where individual queries can be answered.

DSS Data Requests

AUDIENCE

- Those whose question can not be answered through other available data sources

FUNCTION

- Provide group level (aggregate) data or reports

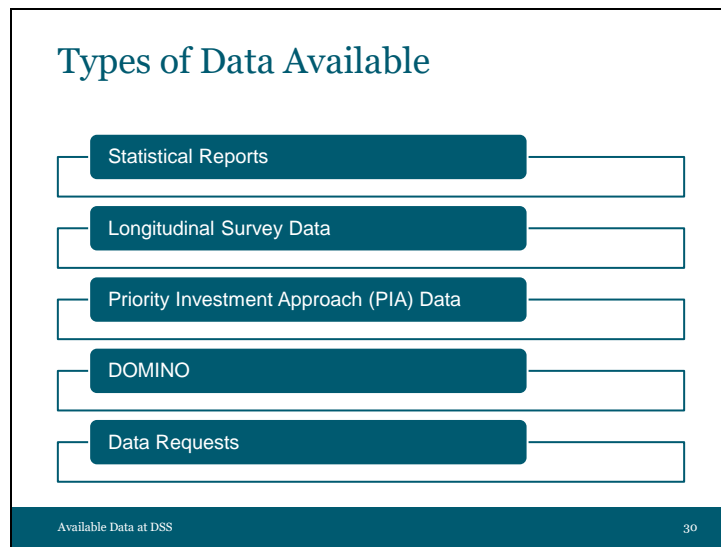
LOCATION

- data.requests@dss.gov.au

Available Data at DSS 29

This service is available for data queries that you haven't been able to address through the other data access points.

In this case, applications can be made for counts of recipients under particular circumstances or conditions. These applications might be with respect to payment data, program data or settlement services. This service is not able to process in-depth research requests but is as a service to provide simple data queries. Please keep in mind, these queries may take time to respond to.



This was a brief (whirlwind) introduction to a range of data services we have at DSS. PIA data in TableBuilder is the most accessible data resource where you are able to seek answers to your specific questions.

The following session will provide a demonstration on how to use TableBuilder, including some of the features and formatting options, using a number of research questions as examples. This session runs for approximately 15-20 minutes.

Any Questions?

Thank you.

Tara Spokes, PhD
Policy Research Officer
tara.spokes@dss.gov.au

Access to Priority Investment Approach (PIA) Data 31

At this point, are there any questions from this session?