

Hi everyone,

Today I will be telling you about a new way to access the Priority Investment Approach research dataset.

In the first session I will provide a short background on the data that is available and briefly introduce one way of accessing the data using TableBuilder. In the second session I will show you a few ways you can use TableBuilder by working through some examples. The second session is for those of you who are interested in a deeper understanding of how to use the dataset. You are welcome to leave after the first session if that doesn't interest you. 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!

We think the Tablebuilder tool will be very useful to a wide number of researchers as well as those applying for TTL Fund grants. I hope that this presentation will enable you to make use of this data yourself.



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

 Both of these examples refer to the Priority Investment Approach data as one source of evidence that you might use. Your evidence is likely to include your own information and data from your experience, however this can be given greater context and supported by the data that is now available to you from the Priority Investment Approach.

So what is the Priority Investment Approach?



What is it?

- The Priority Investment Approach, or PIA, uses data analysis to identify groups at risk of long-term welfare dependence and uses insights from the data to find innovative ways to help more Australians live independently, reducing their need for welfare.
- The PIA data includes administrative data collected for the purpose of recording eligibility for benefits, service delivery activities and payments.



What is in it?

- The data is a series of quarterly snapshots taken from the July-September 2001 quarter to the April-June 2015 quarter.
- The data contains information about those people who have received one of 21 benefits including Aged Pension, Disability Support Pension, Newstart 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.



Where is it?

- In September 2016 the then Minister for Social Services announced a plan to allow researchers limited access to PIA data.
- There are three points of access that differ by level of information provided, as well as cost and ease of access.
- For example, those researchers who have a high level of skill and need more detailed information for their research, can apply for access to the secure enclave (SURE).
- For others, TableBuilder provides non-restricted access to less detailed information with free registration

111100710001			
	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.

This table is a very brief overview of the three access points for the PIA data. It shows the range of access types we have developed, to reflect the range of possible uses and users we expect to be interested in the PIA data. I would like to point out the relative accessibility of the Tablebuilder

- (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 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.





Who it is meant for?

Users are expected to include teachers and students in an educational context, journalists preparing data for reports, as well as academics and NGO's looking for data to use in their grant applications.

• I will provide some examples of both tables and graphs at the end of this session and we will work through how to create and download both tables and graphs in the next session.



As this data is publicly accessible, a number of steps have been taken to ensure privacy is maintained. So let's look at what is NOT in it.

- There are Fewer Variables. There are 38 variables in TableBuilder as opposed to the 60 that are available in the secure environment of SURE.
- There is Less Detail. Some variables have been modified, for example a date might be available as year only in TableBuilder where it is available as month and year in the secure enclave.
- There are Fewer People. TableBuilder uses a 5% sample of the original PIA dataset. 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.



- TableBuilder is good for a wide range of data analysis expertise. Some people here are likely to be analysts or researchers and have experience in working with population data, perhaps even have experience with TableBuilder using other datasets
 - Some people here are likely to have no research experience and not be familiar with working with population data at all.
 - Due to it's user interface, TableBuilder suits both of these populations.
- While the 5% serves to protect individual privacy, it does allow for reportable estimates that can be used in reports and grant applications.



For now, I will show you just a few examples of how you might use TableBuilder.



Everything you can see with TableBuilder is *always* group level or aggregate data. You will never see the individual unit level data that sits 'behind' the tables.

Here, we have a table that displays the number of recipients of two benefit types (Parenting Payment Partnered and Parenting Payment Single)

- Across the top in Columns, and the states and territories,
- Down the left in Rows,
- for June 2015.



- Switching between Table View and Graph View allows you to visualise the data
- Here we can see the count of the two parenting payments for each state.



• By using the options on the left we can now see the same data displayed as the number of recipients across the states for each payment.



We can break this down to see the number of recipients for

- PPP and PPS by
- age groups for each of the states (I have zoomed out so that you can see the whole table).
- We also have the option of showing or hiding totals. In this case we can see the total count for the country as well as the count for each state.



Switching to Graph View I now have the states across the bottom and can see the count of PPP and PPS recipients by the age groups.

We have the colour key across the bottom for each age group for the two payment types.



Alternatively we can view the data by payment for each age group and comparing the states next to each other.

You can see now we have the colour key across the bottom for each state.



- In this table I have chosen to compare recipient's work capacity prior to an intervention and after an intervention. That is, the number of hours per week a recipient is assessed as being able to work. These will include those recipients who have partial working capacity and have had a Job Capacity Assessment or Employment Services Assessment where there are barriers to work.
- In this case, where a recipient was assessed as having 8-14 hours capacity prior to an intervention and assessed as having the same 8-14 hours capacity with an intervention, they are represented here
- Where a recipient was assessed as having 8-14 hours capacity prior to an intervention and then assessed as having 15-22 hours capacity with an intervention, they are represented here



- Using a different graphing option we can visually represent the trend of an increasing number of hours assessed for work capacity with an intervention.
- For example you can see that of those who were assessed as being able to work 15-22 hours after an intervention, a large proportion were able to work 8-14 hours prior to an intervention (dark green).
- Similarly, a large proportion of those who were assessed as being able to work 23-29 hours after an intervention were able to work 15-22 hours prior to an intervention (light green).

That is we can see an increase in capacity as assessed from a Work Capacity Assessment after an intervention.



Now to look at changes in primary medical condition across time, this table presents the primary medical condition in the rows and the April-June quarters across four different years across the top in the columns.



We can see here that both

- musculoskeletal conditions and psychological or psychiatric conditions are the two main primary medical conditions. In this case we are looking at
- people receiving Disability Support Pension. By comparing across the four years, we can also see that this number is increasing
- for psychological or psychiatric conditions.



We can use

- this drop box to compare medical condition to recipients of other benefit types, in this case we are looking at those receiving Newstart Allowance. In this case, the number of recipients with psychological or psychiatric condition as their primary medical condition is
- decreasing across the years.



By looking at all recipients, regardless of the type of benefit

• (using 'total' figures), we can see that **overall** the number of people whose primary medical condition is psychological or psychiatric in nature, is increasing across the years.



The following session will provide some basic instructions 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. I also have a list of the available variables and categories in the dataset for you to refer to in the break if you like. Are there any questions from this session?