Unlocking the potential of the Integrated Data Infrastructure for research

RezBaz July 2025

Social Investment Agency
Toi Hau Tāngata

Te Kāwanatanga o Aotearoa
New Zealand Government

Insight arising from integration

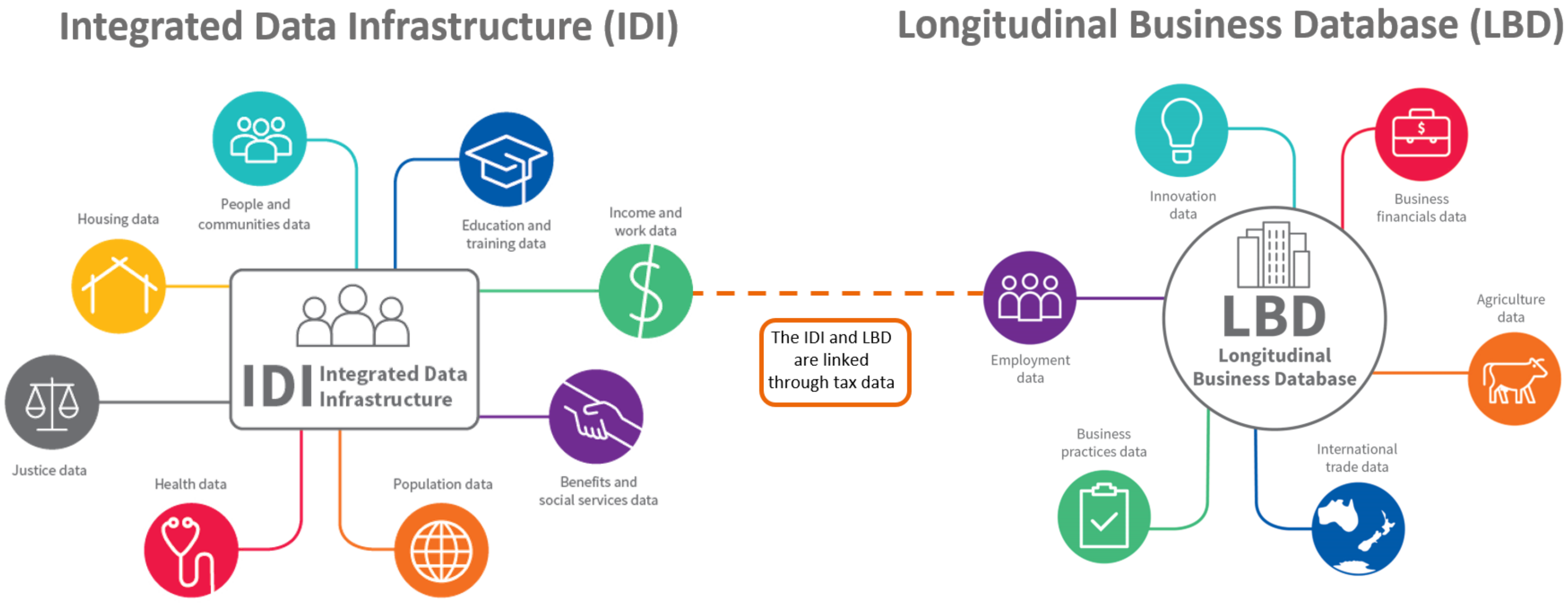
Many insights arise from bringing together information that was once separate

Integrated data is designed for this purpose

- What is the Integrated Data Infrastructure (IDI)
  - What is available and how it is integrated
  - Data protections in place
- What has been done with it
  - Examples of projects
  - How integration enables questions to be answered
- Some tips, tricks, and resources for using it
  - Significant data wrangling with steep learning curve
  - Range of resources to assist new researchers

What is the integrated data infrastructure?

World leading tool for research and analysis


Integrated Data Infrastructure (IDI)
Longitudinal Business Database (LBD)
Housing data
People and communities data
Education and training data
Income and work data
IDI Integrated Data Infrastructure
Justice data
Health data
Population data
Benefits and social services data
Innovation data
Business financials data
Employment data
LBD Longitudinal Business Database
Agriculture data
Business practices data
International trade data

The IDI and LBD are linked through tax data

The power of integrated data

Combine information

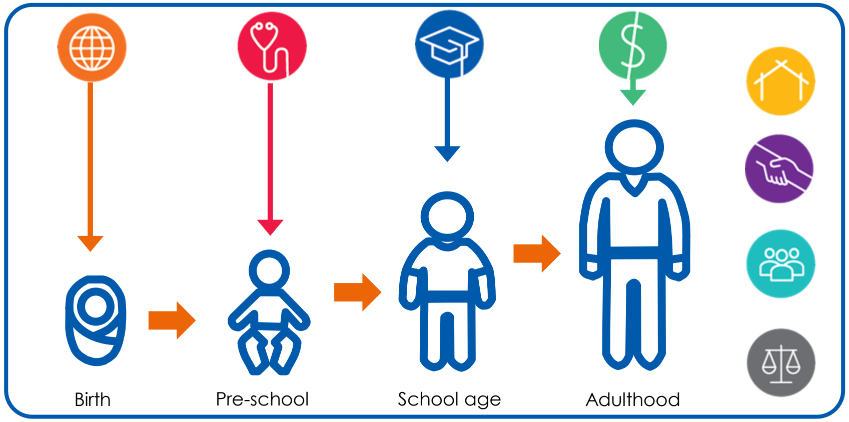
- from multiple domains
- from multiple life stages

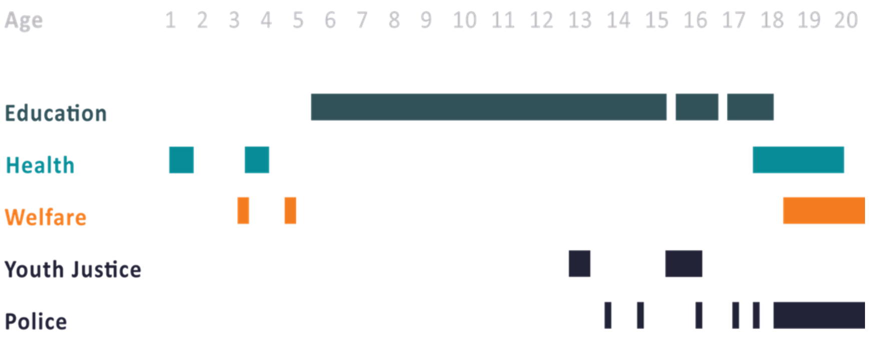
Construct detailed cross-sections

Follow experiences over time

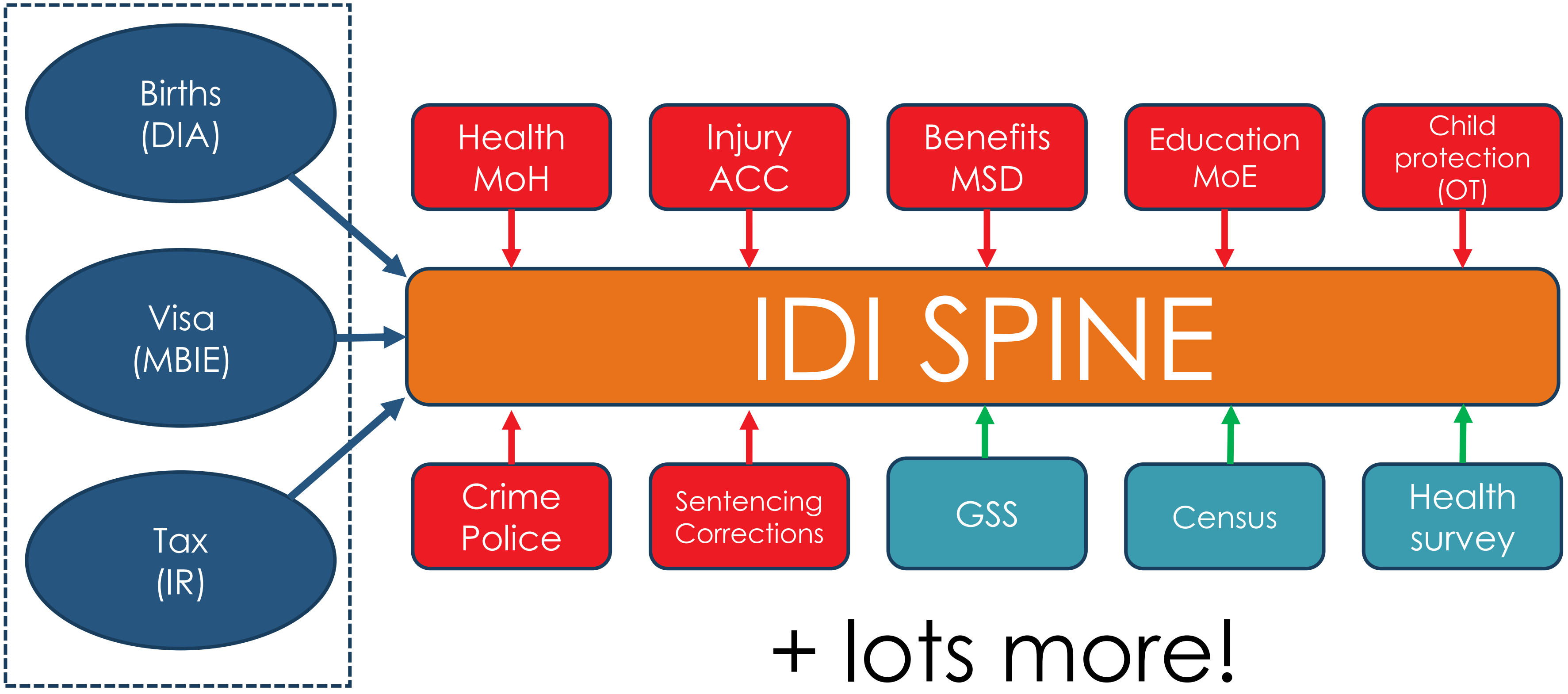
- Interactions between services
- Panel and longitudinal data

Predict or track outcomes



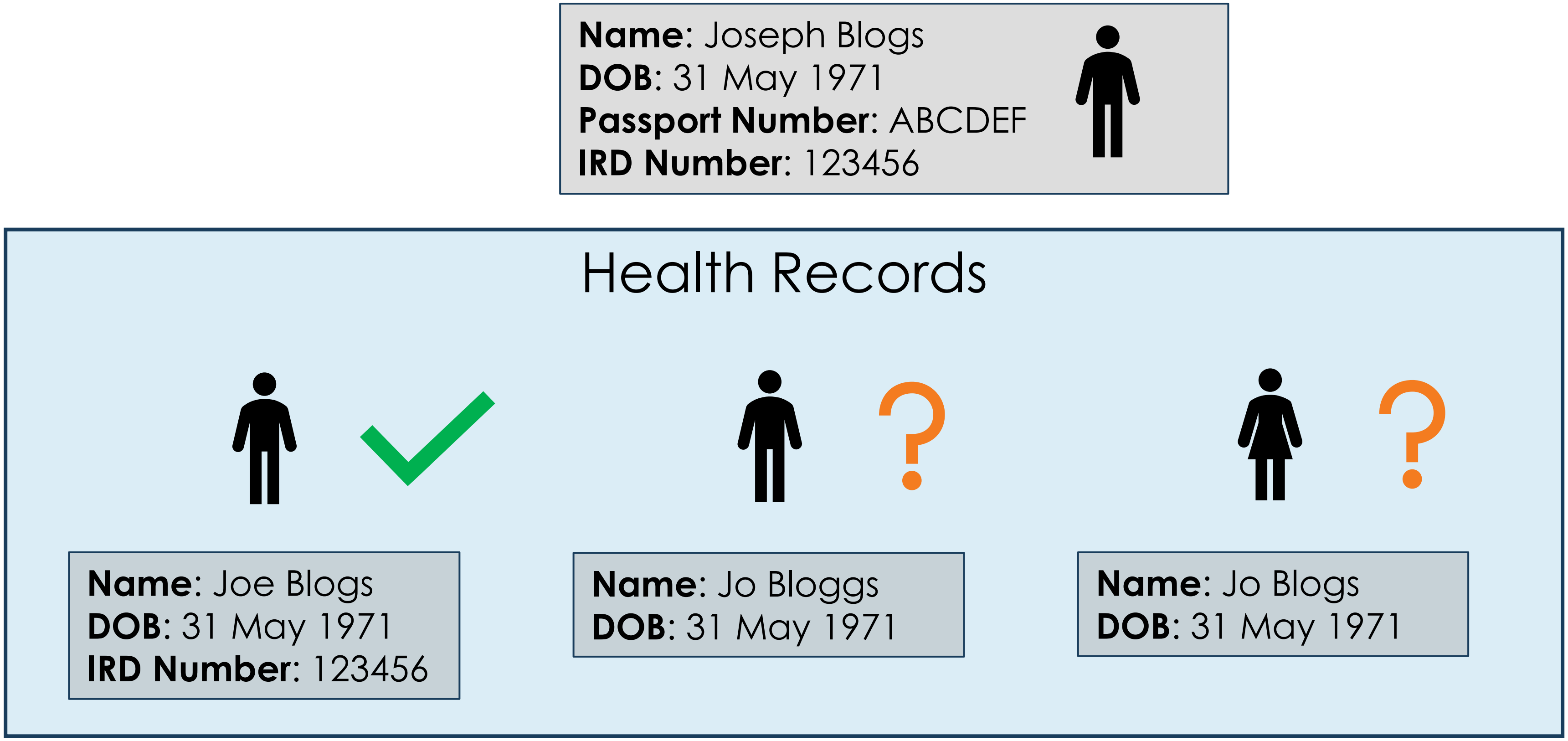


A list of all identities – the IDI Spine



+ lots more!

Deterministic vs Probabilistic linking



Name: Joseph Blogs
DOB: 31 May 1971
Passport Number: ABCDEF
IRD Number: 123456

Health Records

✓

Name: Joe Blogs
DOB: 31 May 1971
IRD Number: 123456

?

Name: Jo Bloggs
DOB: 31 May 1971

?

Name: Jo Blogs
DOB: 31 May 1971

IDI Data sets vary in history and currency

Pre-1985

2000

2010

2024

Health

Education and Training

Benefits and Social Services

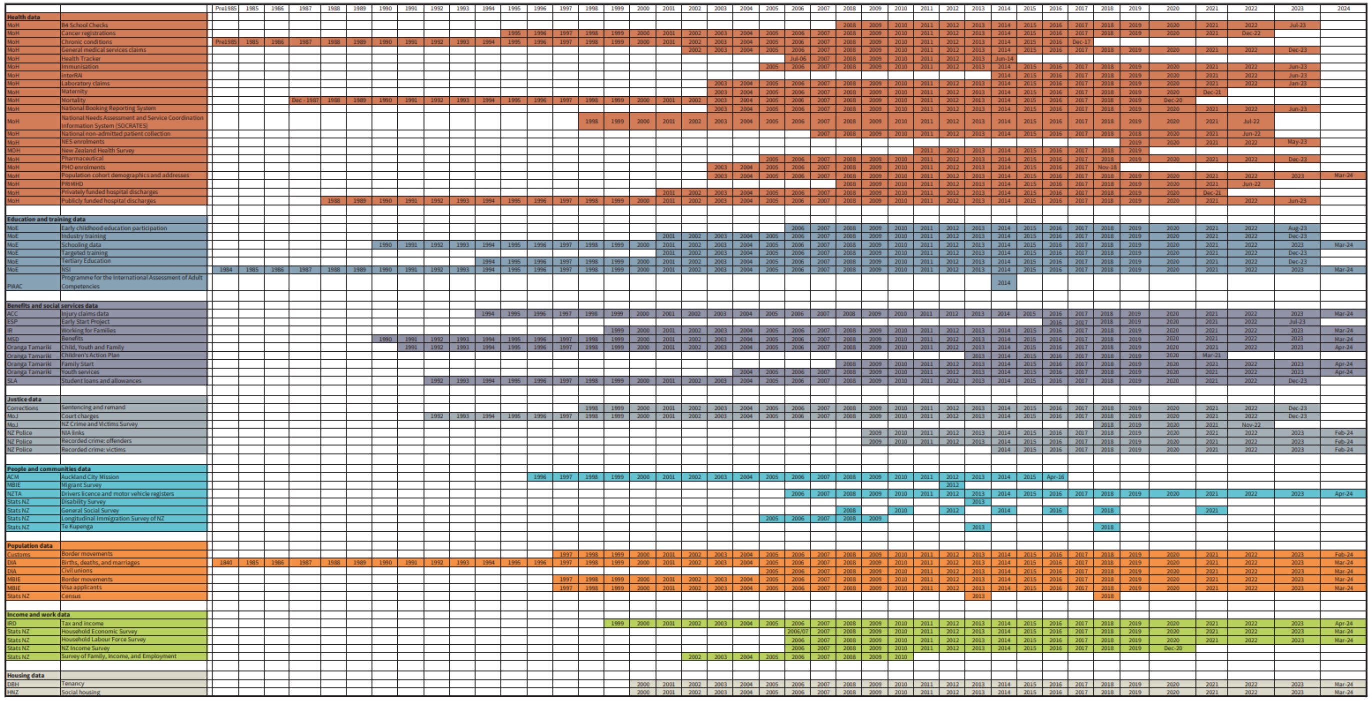
Justice

People and communities

Population Data

Income and Work

Housing Data



		Pre1985	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024
Health data																																										
MoH	B4 School Checks																									2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	Jul-23	
MoH	Cancer registrations												1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	Dec-22		
MoH	Chronic conditions	Pre1985	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	Dec-17							
MoH	General medical services claims																			2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	Dec-23	
MoH	Health Tracker																							Jul-06	2007	2008	2009	2010	2011	2012	2013	Jun-14										
MoH	Immunisation																						2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	Jun-23	
MoH	interRAI																															2014	2015	2016	2017	2018	2019	2020	2021	2022	Jun-23	
MoH	Laboratory claims																				2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	Jan-23	
MoH	Maternity																				2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	Dec-21			
MoH	Mortality				Dec - 1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	Dec-20				
MoH	National Booking Reporting System																				2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	Jun-23	
MoH	National Needs Assessment and Service Coordination Information System (SOCRATES)															1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	Jul-22		
MoH	National non-admitted patient collection																								2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2018	2020	2021	Jun-22		
MoH	NES enrolments																																				2019	2020	2021	2022	May-23	
MOH	New Zealand Health Survey																												2011	2012	2013	2014	2015	2016	2017	2018	2019					
MoH	Pharmaceutical																						2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	Dec-23	
MoH	PHO enrolments																				2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	Nov-18						
MoH	Population cohort demographics and addresses																				2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	Mar-24
MoH	PRIMHD																									2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	Jun-22		
MoH	Privately funded hospital discharges																		2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	Dec-21			
MoH	Publicly funded hospital discharges					1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	Jun-23	
Education and training data																																										
MoE	Early childhood education participation																							2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	Aug-23	
MoE	Industry training																		2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	Dec-23	
MoE	Schooling data							1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	Mar-24
MoE	Targeted training																		2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	Dec-23	
MoE	Tertiary Education											1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	Dec-23	
MoE	NSI	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	Mar-24
PIAAC	Programme for the International Assessment of Adult Competencies																															2014										
Benefits and social services data																																										
ACC	Injury claims data											1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	Mar-24
ESP	Early Start Project																																	2016	2017	2018	2019	2020	2021	2022	Jul-23	
IR	Working for Families																1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	Mar-24
MSD	Benefits							1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	Mar-24
Oranga Tamariki	Child, Youth and Family								1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	Apr-24
Oranga Tamariki	Children's Action Plan																														2013	2014	2015	2016	2017	2018	2019	2020	Mar-21			
Oranga Tamariki	Family Start																									2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	Apr-24
Oranga Tamariki	Youth services																					2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	Apr-24
SLA	Student loans and allowances									1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	Dec-23	
Justice data																																										
Corrections	Sentencing and remand															1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	Dec-23	
MoJ	Court charges									1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	Dec-23	
MoJ	NZ Crime and Victims Survey																																			2018	2019	2020	2021	Nov-22		
NZ Police	NIA links																										2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	Feb-24
NZ Police	Recorded crime: offenders																										2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	Feb-24
NZ Police	Recorded crime: victims																															2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	Feb-24
People and communities data																																										
ACM	Auckland City Mission													1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	Apr-16								
MBIE	Migrant Survey																													2012												
NZTA	Drivers licence and motor vehicle registers																							2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	Apr-24
Stats NZ	Disability Survey																														2013											
Stats NZ	General Social Survey																									2008		2010		2012		2014		2016		2018			2021			
Stats NZ	Longitudinal Immigration Survey of NZ																						2005	2006	2007	2008	2009															
Stats NZ	Te Kupenga																														2013					2018						
Population data																																										
Customs	Border movements														1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	Feb-24
DIA	Births, deaths, and marriages	1840	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	Mar-24
DIA	Civil unions																						2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	Mar-24
MBIE	Border movements														1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	Mar-24
MBIE	Visa applicants														1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	Mar-24
Stats NZ	Census																														2013					2018						
Income and work data																																										
IRD	Tax and income																1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	Apr-24
Stats NZ	Household Economic Survey																							2006/07	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	Mar-24
Stats NZ	Household Labour Force Survey																							2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	Mar-24
Stats NZ	NZ Income Survey																							2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	Dec-20				
Stats NZ	Survey of Family, Income, and Employment																			2002	2003	2004	2005	2006	2007	2008	2009	2010														
Housing data																																										
DBH	Tenancy																	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	Mar-24
HNZ	Social housing																	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	Mar-24

‘Five Safes’ keep integrated data safe

Safe People

Only approved researchers can access or view microdata.

Safe Projects

Data can only be used for research projects in the public interest.

Safe Settings

Research takes place in secure data labs which Stats controls.

Safe Data

Access is granted only to the data that is needed for the research.

Safe Output

Confidentiality rules protect against privacy breaches.
All output is checked by Stats NZ to confirm it is safe.

Data is de-identified

Information supplied to Stats NZ

- **Name:** Star Thinker
- **Date of birth:** 29 February 1933
- **IRD:** 123-123-123
- **NHI:** 0123456789
- **Address:** 123 Enlightenment Terrace, Researchville

Information visible to researchers

- **snz_uid:** 4545454545
- **Birth month:** May
- **Birth year:** 1981
- **snz_ird_uid:** 111111
- **snz_moh_uid:** 22222
- **address_uid:** 99999
- **Meshblock:** 4507

Confidentiality rules limit data release

Microdata output guide specifies output rules

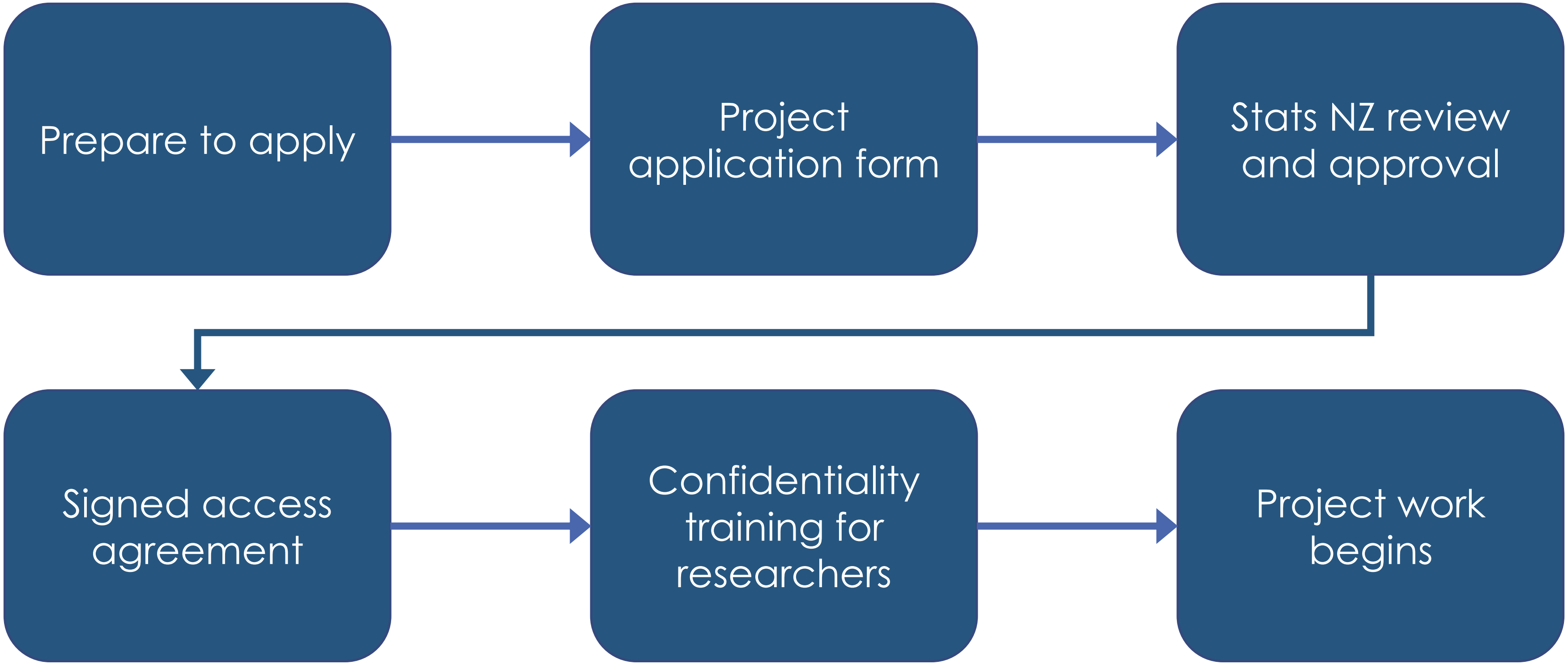
Four most common rules:

- Random rounding of counts to base 3
- Counts of fewer than 6 people are suppressed
- Totals for fewer than 20 people are suppressed
- Values that reflect a single organization are suppressed

Example process

Output	Raw	Released
People in Researchville	62	60
Total income for people in Researchville	$999999	$999999
Academics in Researchville	17	18
Total income for academics in Researchville	$777777	Suppress
Award winners in Researchville	5	Suppress
Total income for award winners in Researchville	$555555	Suppress
Employees of Top Research Inc	8	Suppress

Applying for access



Prepare to apply

Project application form

Stats NZ review and approval

Signed access agreement

Confidentiality training for researchers

Project work begins

Apply to use microdata for research
www.stats.govt.nz/integrated-data/apply-to-use-microdata-for-research/

Not perfect – has strengths and limitations

Rich range of data

Skilled and in-demand roles

Secure access

Enforced privacy protections

Range of data quality and documentation shortcomings

Technical capability and knowledge requirements

Administrative process
Barriers to collaboration

Small number studies limited

What has been done with integrated data?

Types of questions the IDI is good at

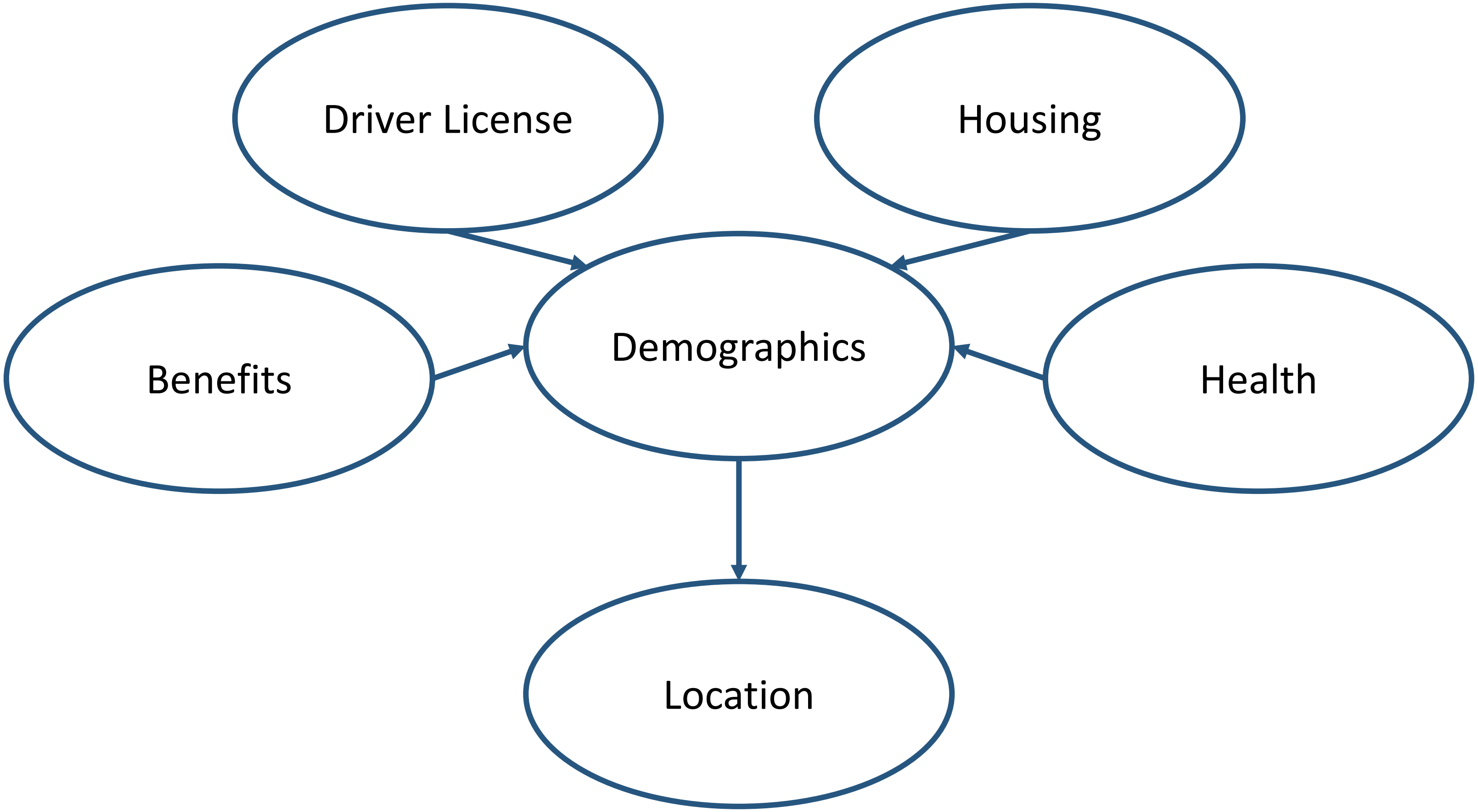
Descriptive

Inferential

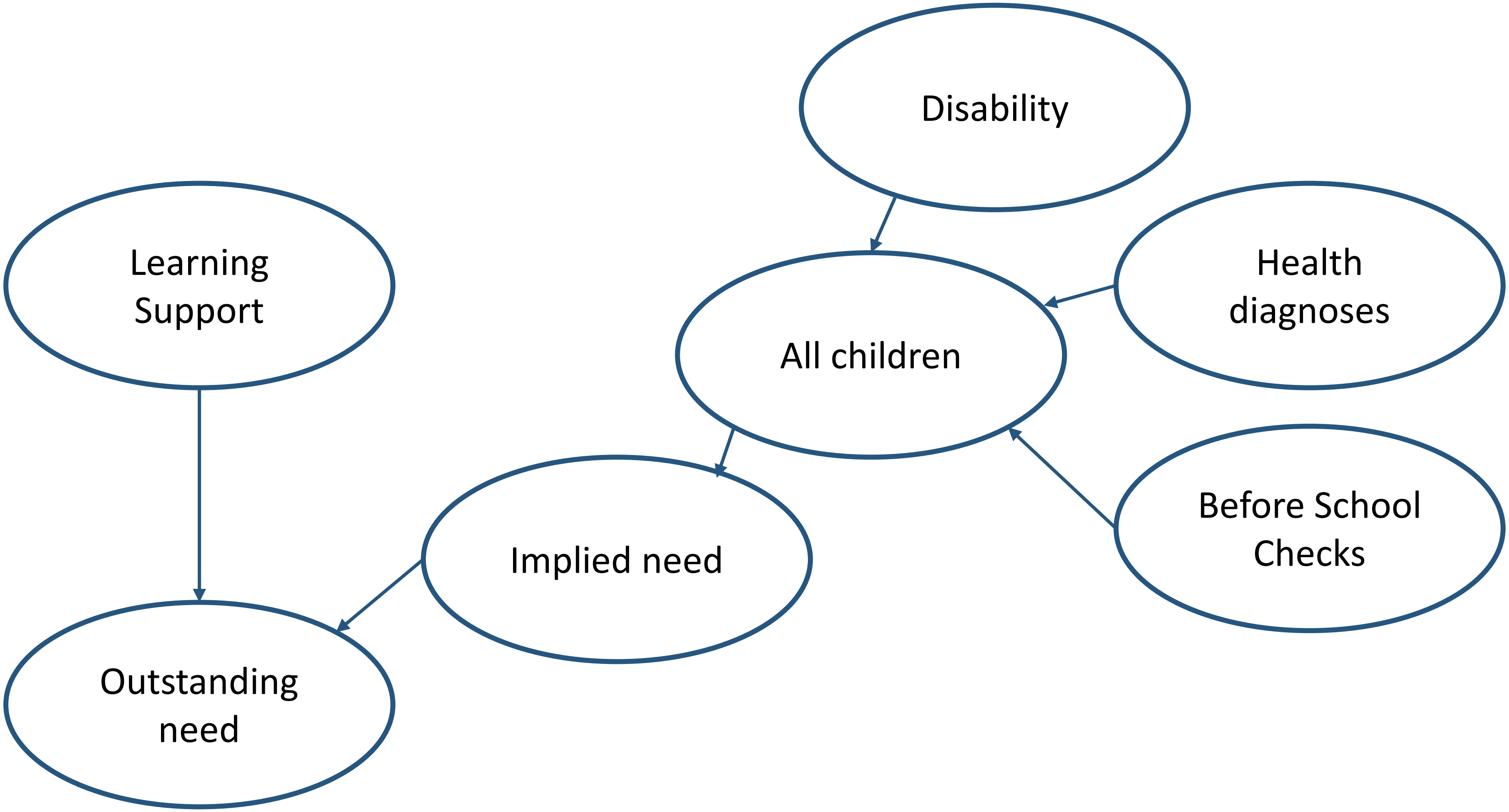
Predictive

Overlaps	Unmet needs	Impact	Lifecourse	Simulation
What don't I know about my clients? Finding people who have an interaction with agency A (or other characteristic) and looking at what other agencies know about them.	*'Who else should we provide services to?'* The inverse of overlaps – who are the people who we don't see in agency A, but they look like agency A clients?	*'Are we making a difference?'* The richness of data in the IDI can help in identifying comparison groups; the longitudinal nature allows for follow-up.	'What leads to the outcomes we want (or don't want)?' Longitudinal analysis that could be defined by an experience/ outcome at the start, middle or end of period.	*'Where should we take action? What would be the likely impact?'* Many models exist in the IDI that can be used for forecasting or 'what if' analysis.

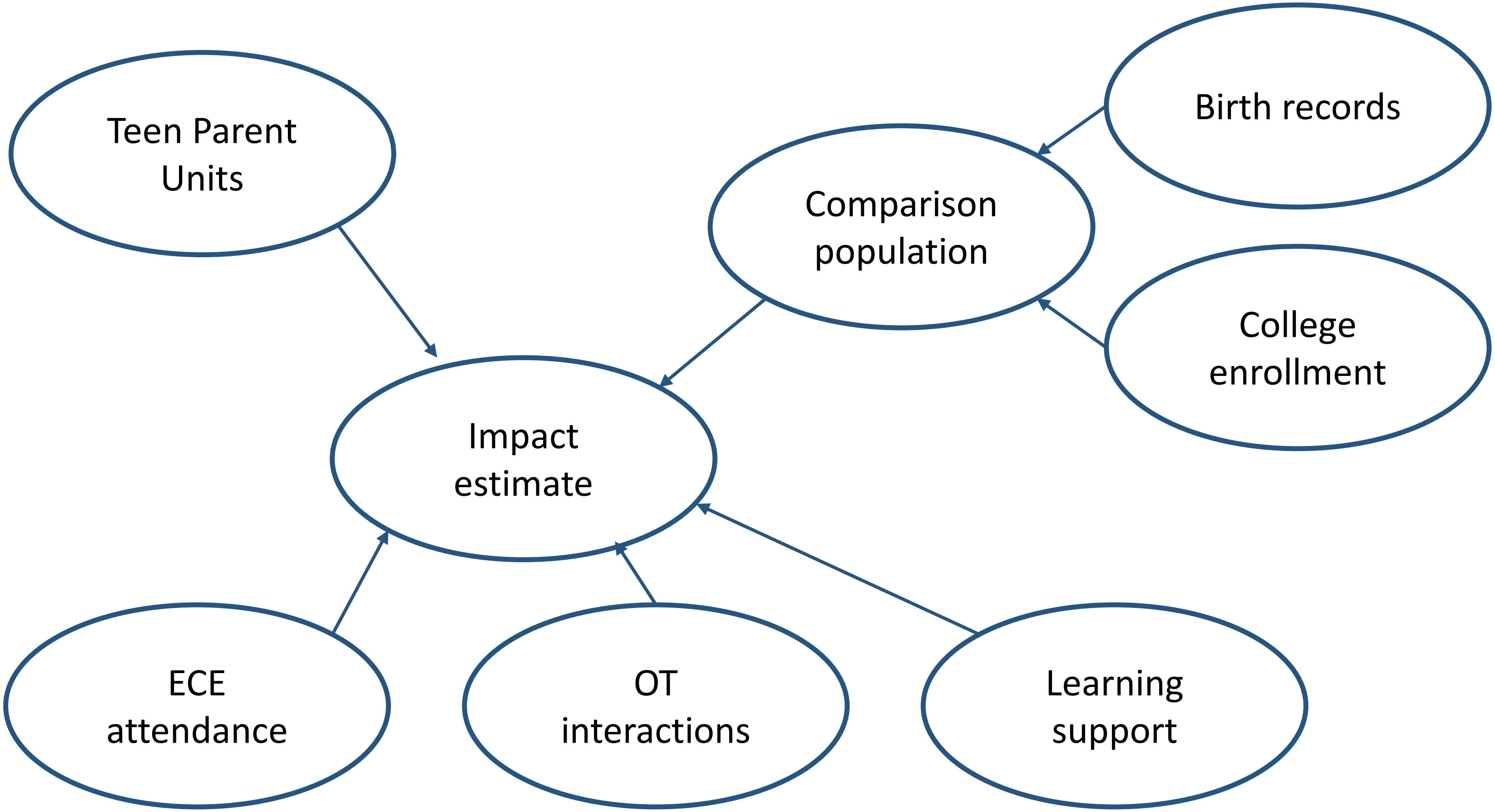
Older people experiencing vulnerablity



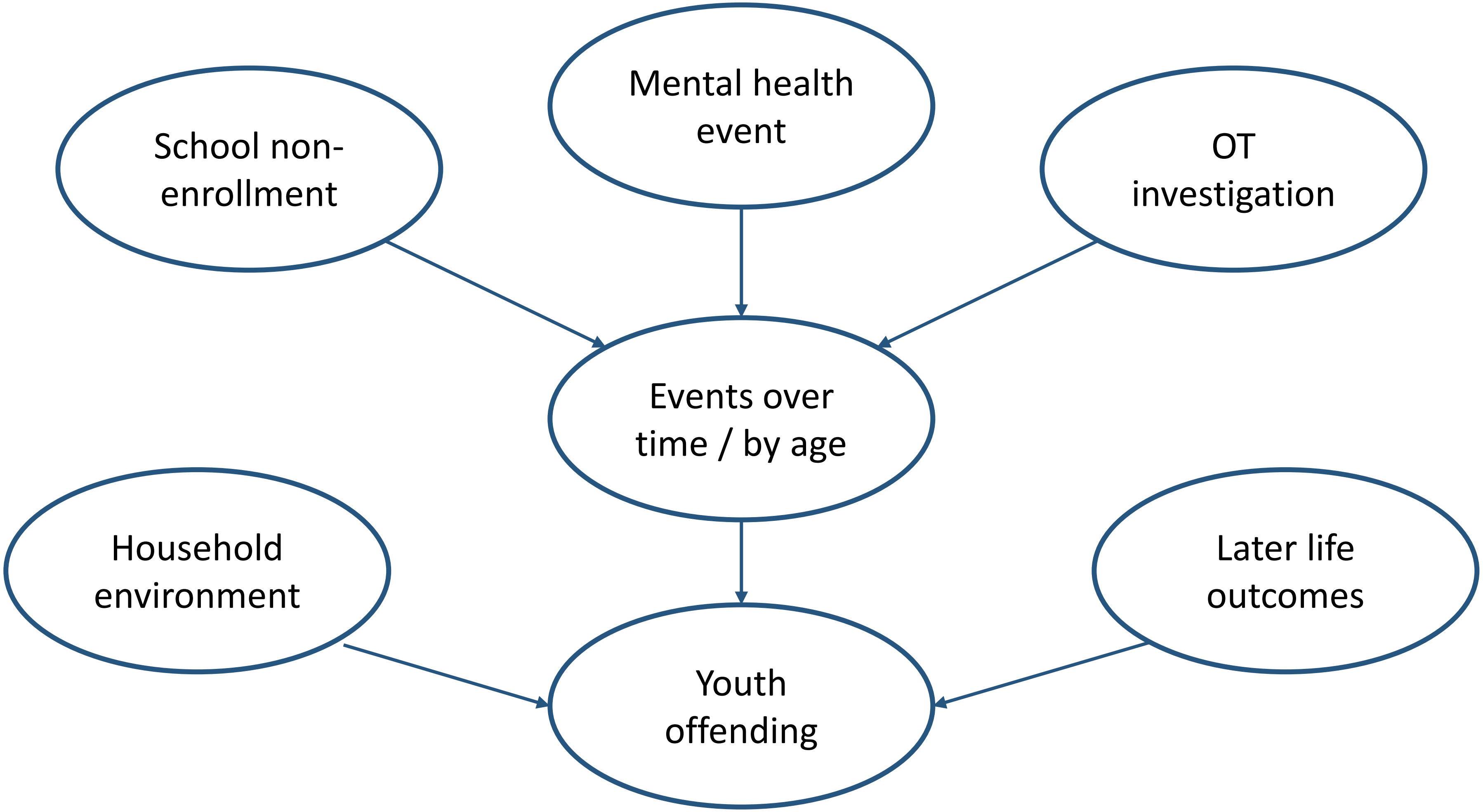
Highest learning needs review



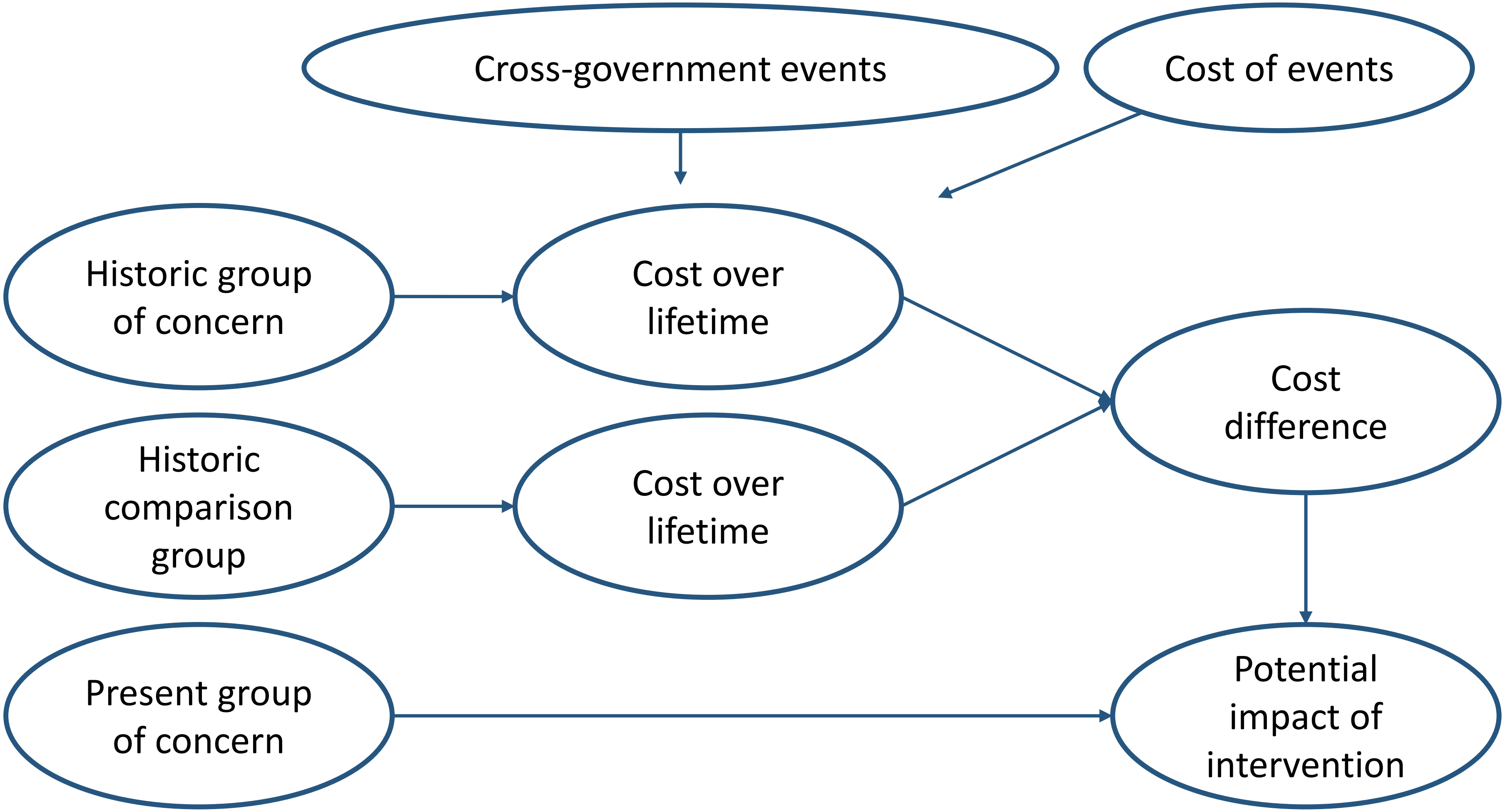
Long-term impact of teen parent units



Journeys of youth offenders



Changes in trajectory



Cross-government events

Cost of events

Historic group of concern

Cost over lifetime

Cost difference

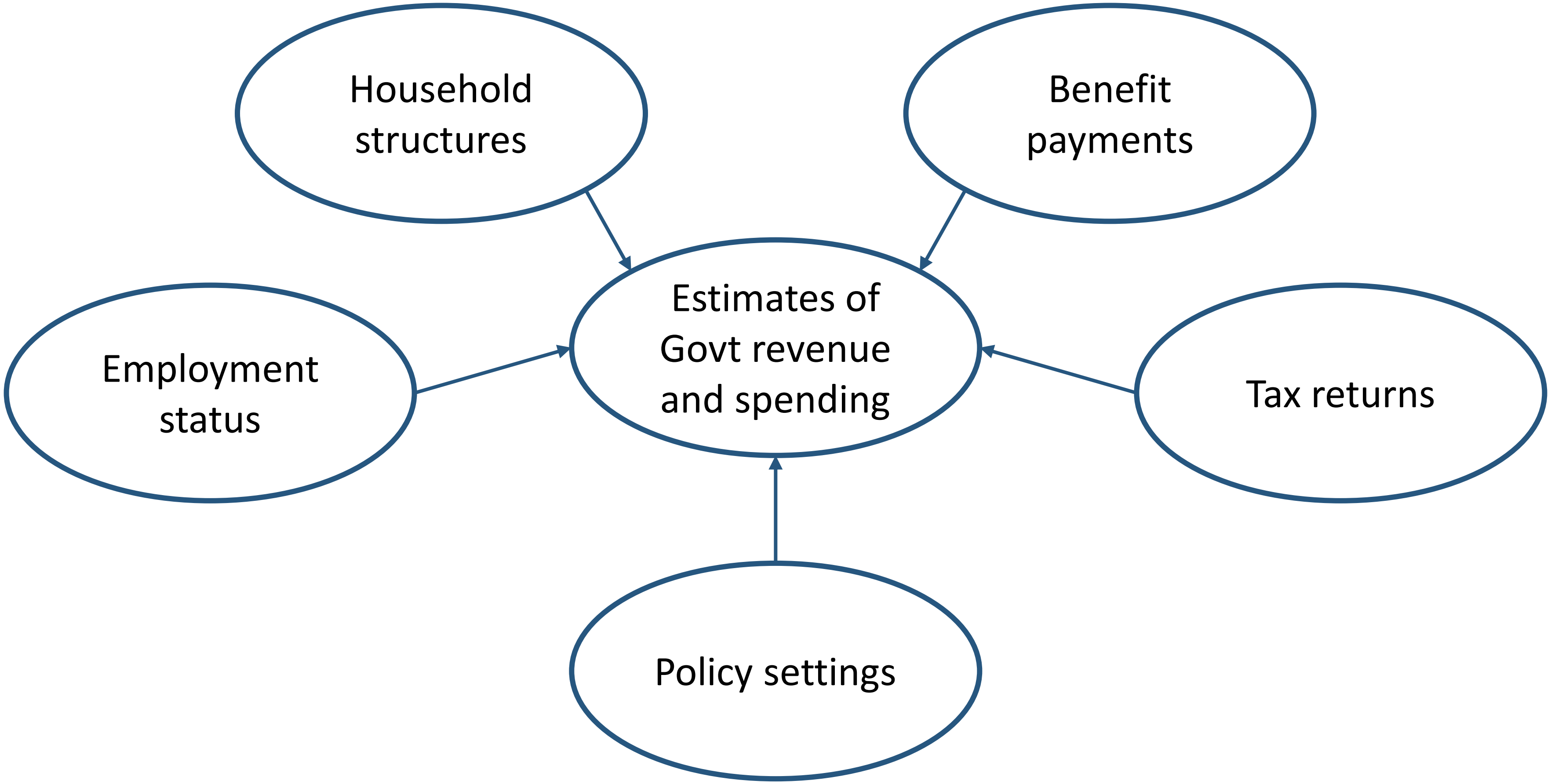
Historic comparison group

Cost over lifetime

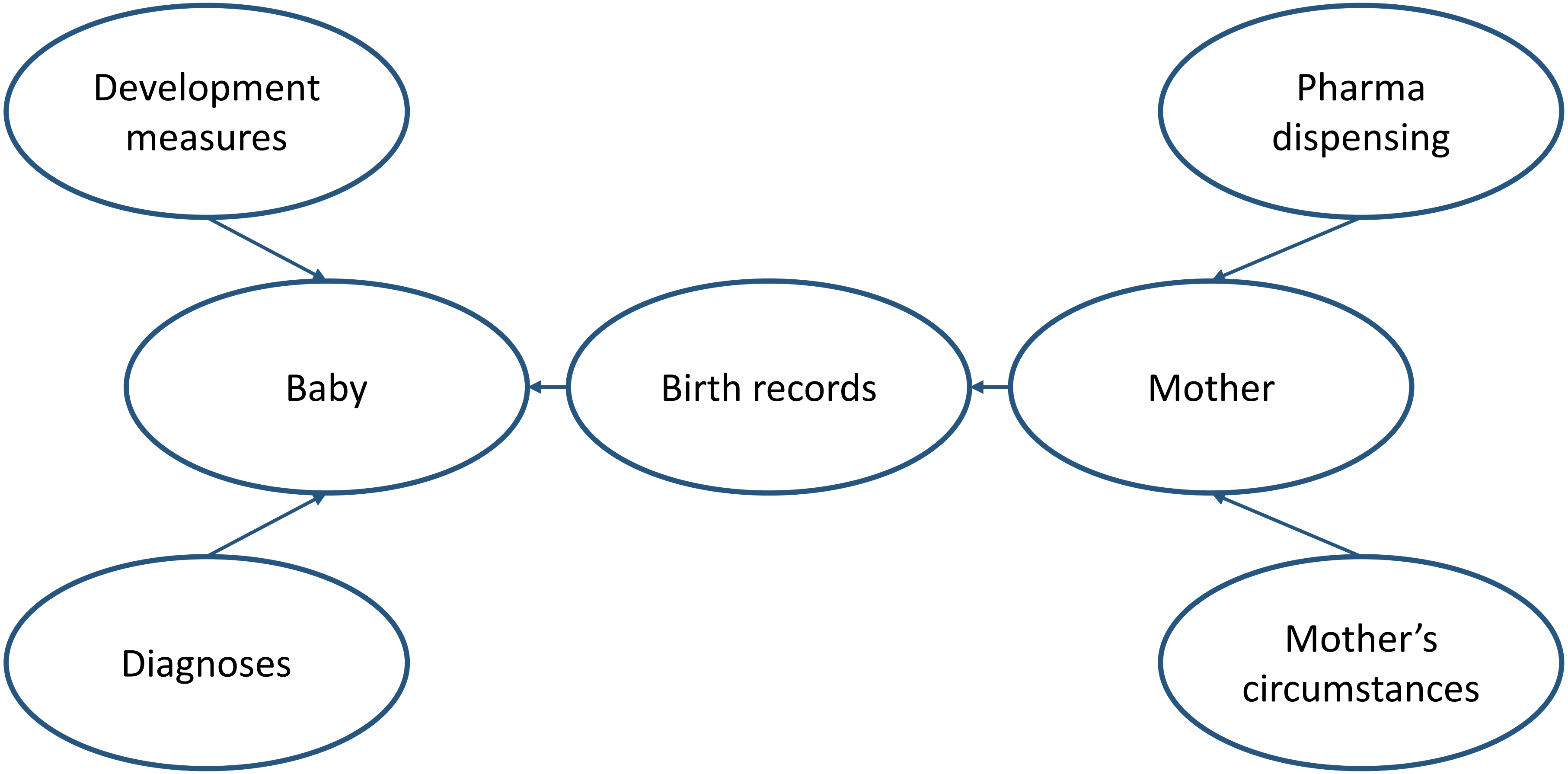
Present group of concern

Potential impact of intervention

Tax policy microsimulation modelling



Effect of maternal antibiotic use of babies



Development measures

Pharma dispensing

Baby

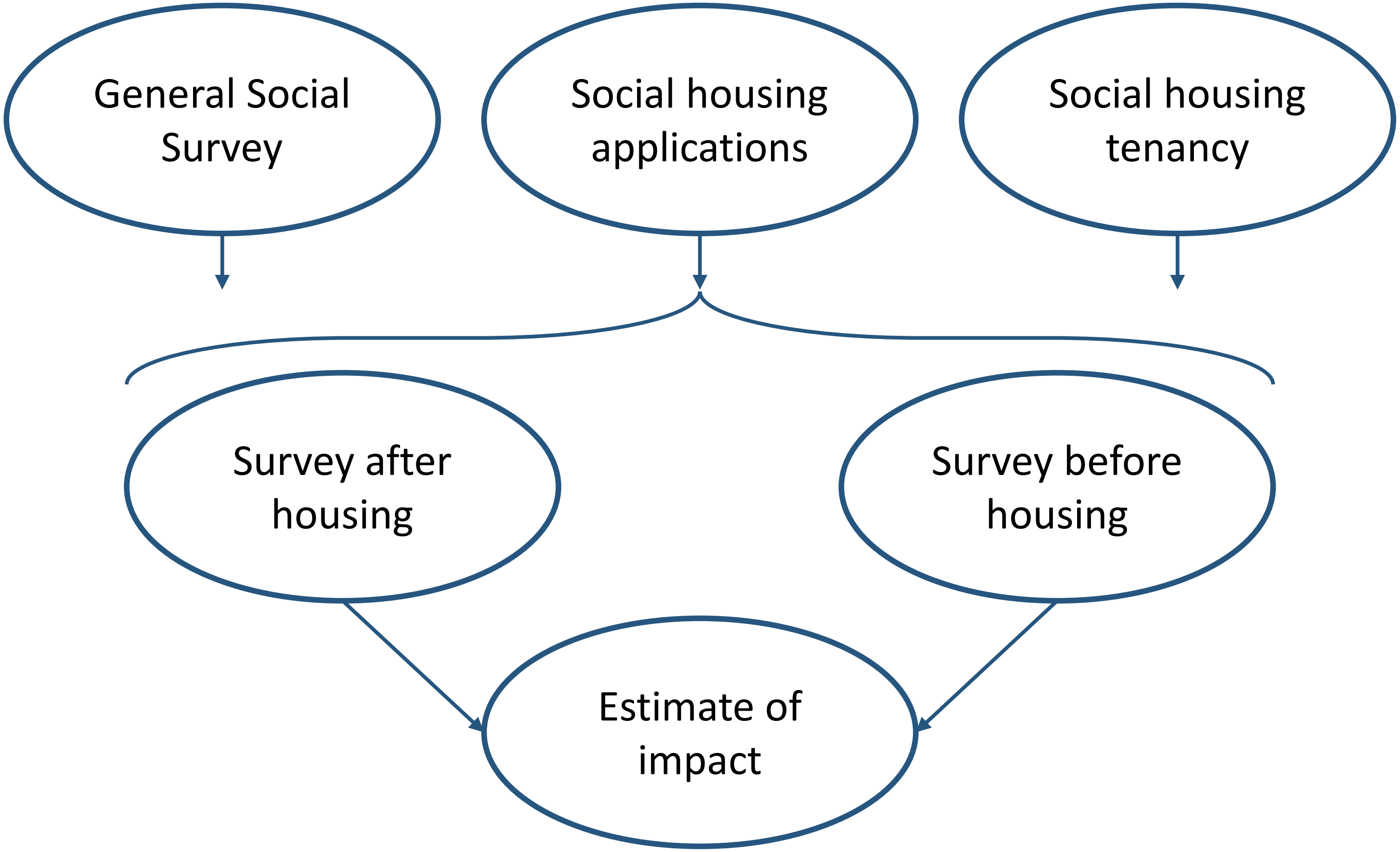
Birth records

Mother

Diagnoses

Mother’s circumstances

Wellbeing impact of social housing


General Social Survey

Social housing applications

Social housing tenancy

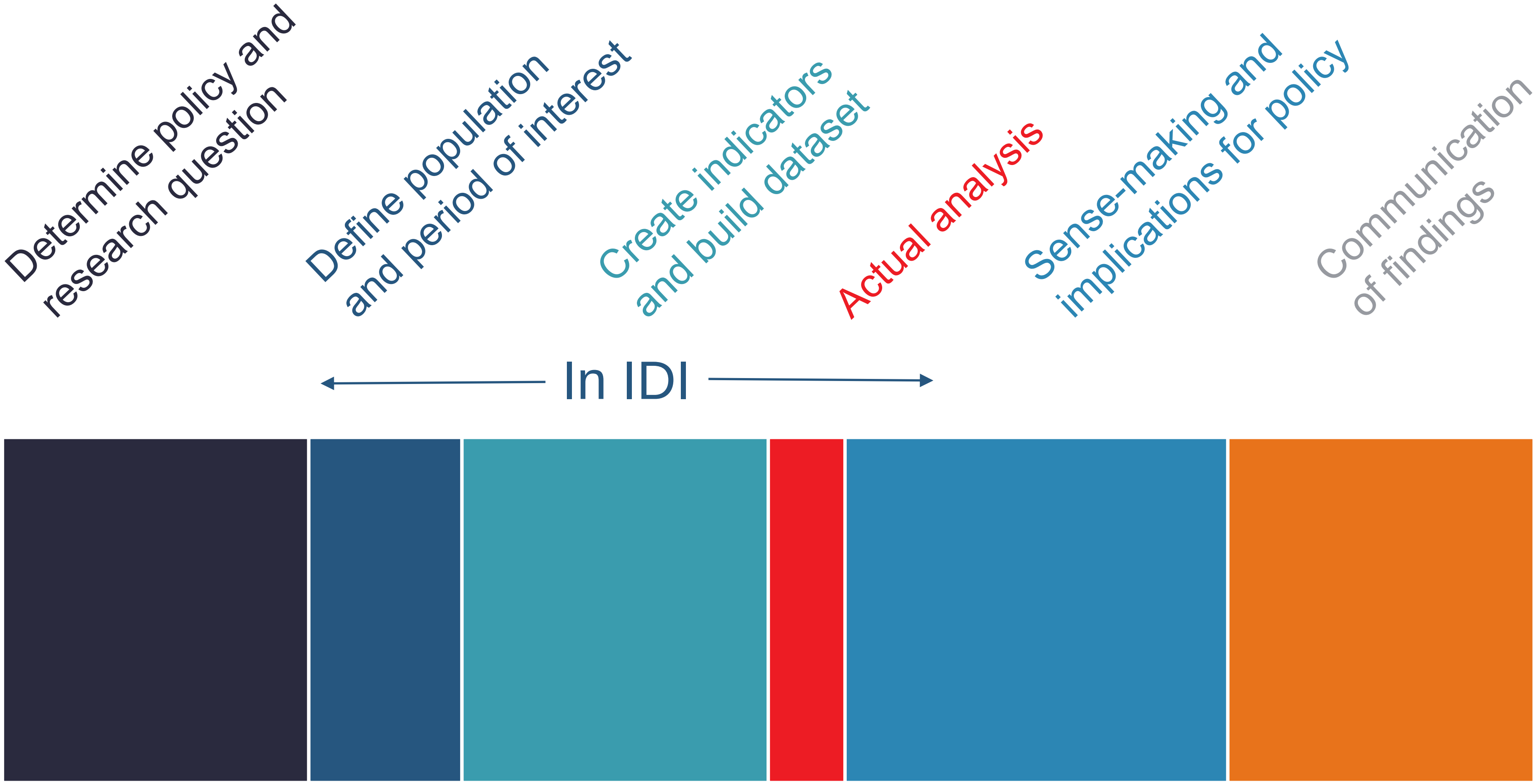
Survey after housing

Survey before housing

Estimate of impact

What advice makes it easier to use integrated data?

Lifecycle of an IDI project (our experience)



Determine policy and research question

Define population and period of interest

Create indicators and build dataset

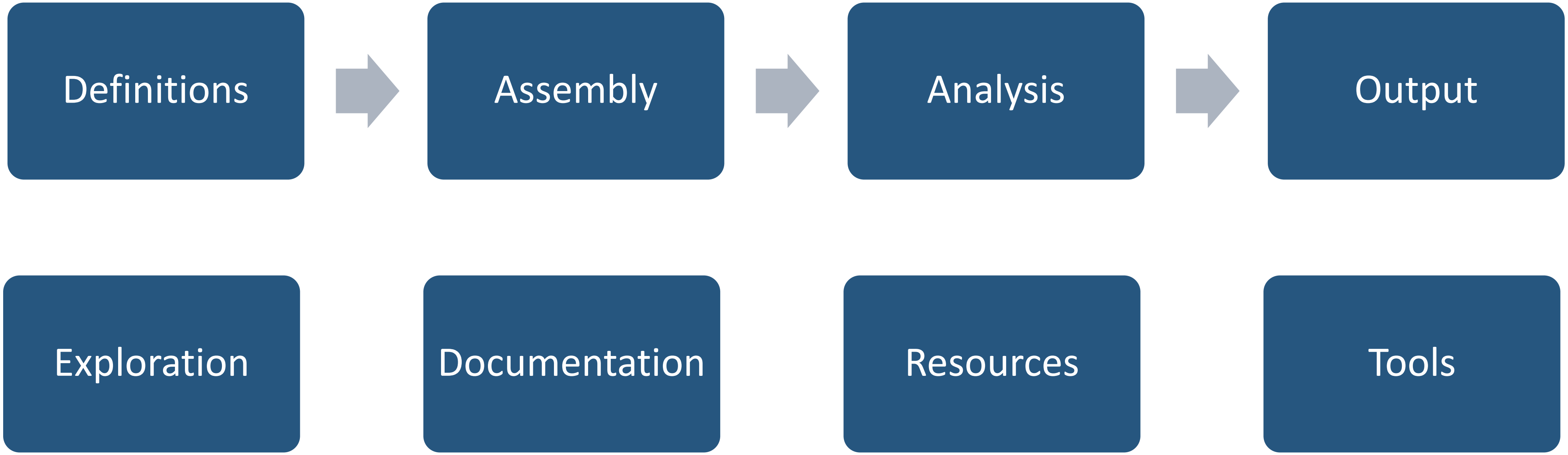
Actual analysis

Sense-making and implications for policy

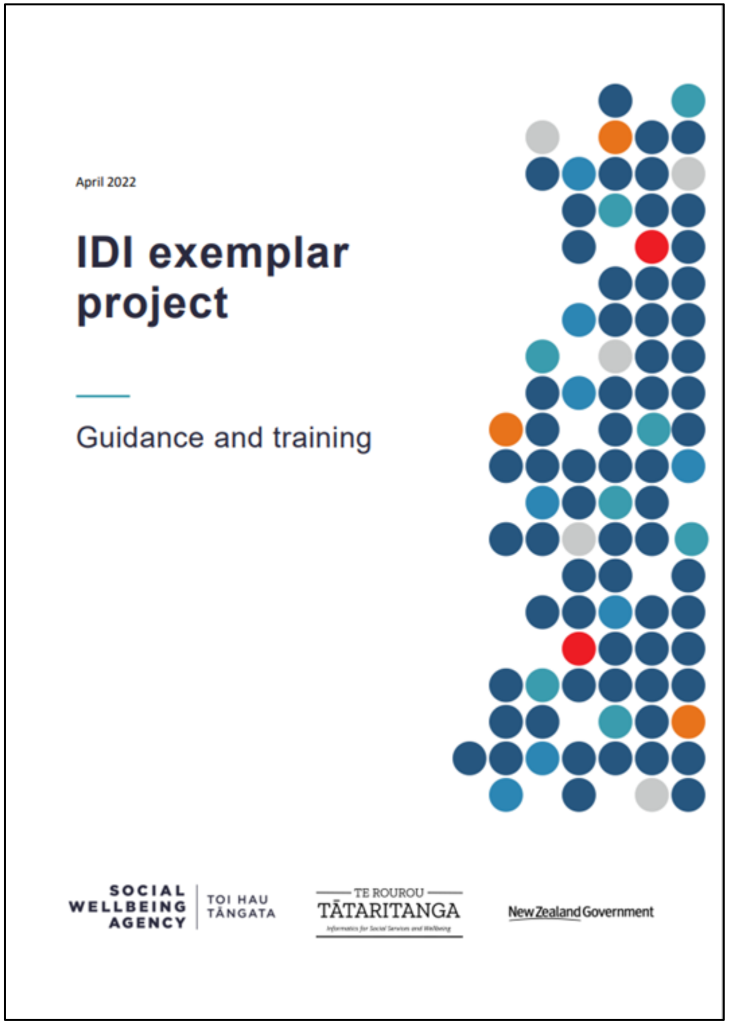
Communication of findings

In IDI

Structure projects using consistent patterns



Start with a small project

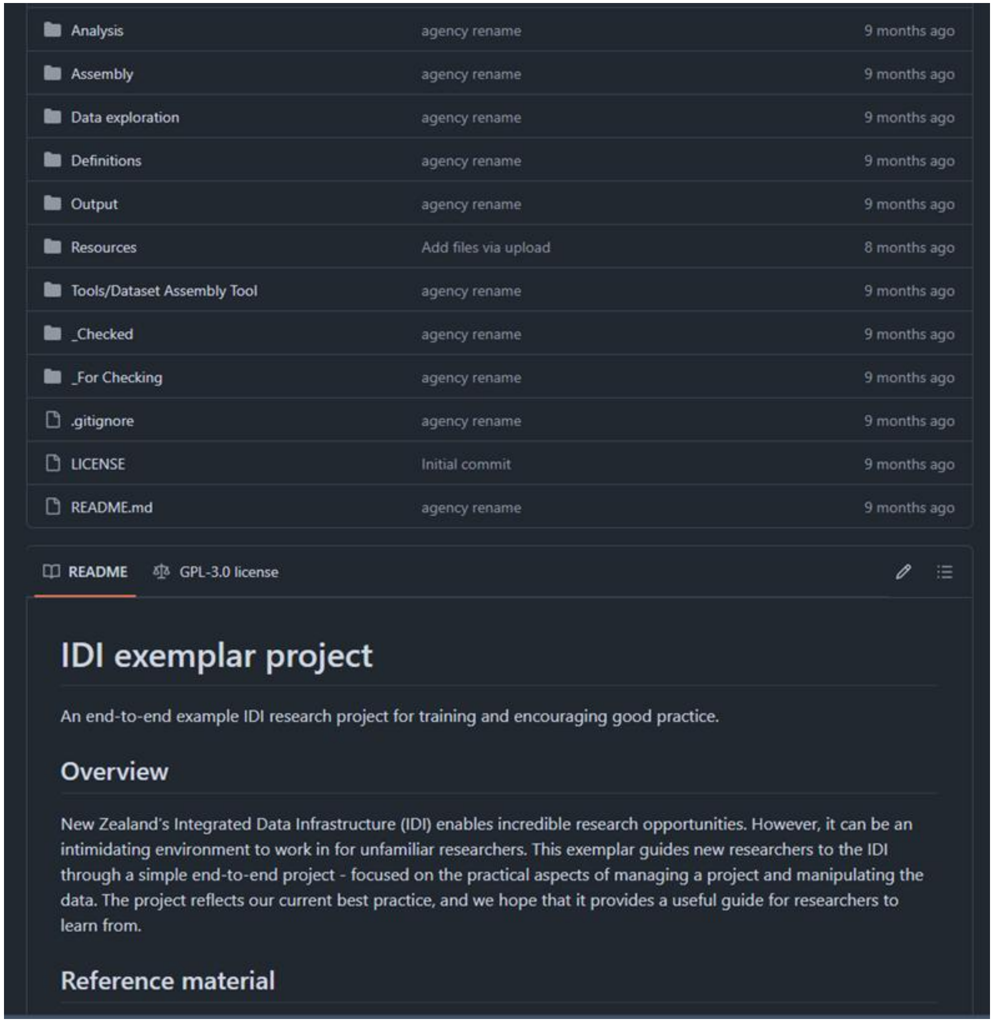


April 2022

IDI exemplar project

Guidance and training

SOCIAL WELLBEING AGENCY | TOI HAU TĀNGATA
TE ROUROU TĀTARITANGA
New Zealand Government



Analysis	agency rename	9 months ago
Assembly	agency rename	9 months ago
Data exploration	agency rename	9 months ago
Definitions	agency rename	9 months ago
Output	agency rename	9 months ago
Resources	Add files via upload	8 months ago
Tools/Dataset Assembly Tool	agency rename	9 months ago
_Checked	agency rename	9 months ago
_For Checking	agency rename	9 months ago
.gitignore	agency rename	9 months ago
LICENSE	Initial commit	9 months ago
README.md	agency rename	9 months ago

IDI exemplar project

An end-to-end example IDI research project for training and encouraging good practice.

Overview

New Zealand's Integrated Data Infrastructure (IDI) enables incredible research opportunities. However, it can be an intimidating environment to work in for unfamiliar researchers. This exemplar guides new researchers to the IDI through a simple end-to-end project - focused on the practical aspects of managing a project and manipulating the data. The project reflects our current best practice, and we hope that it provides a useful guide for researchers to learn from.

Reference material

Recalibrate expectations

Example task 1:

Compare self-reported life satisfaction against every other measure in the General Social Survey (GSS).

Non-technical perspective:

Concern that large number of crosstabs will be time consuming to create.

Analytic approach:

Quick and straightforward.

Only one input table, already arranged for analysis. Repetitive processing done by computer not researcher.

Example task 2:

Count the number of benefit recipients with children who have diabetes.

Non-technical perspective:

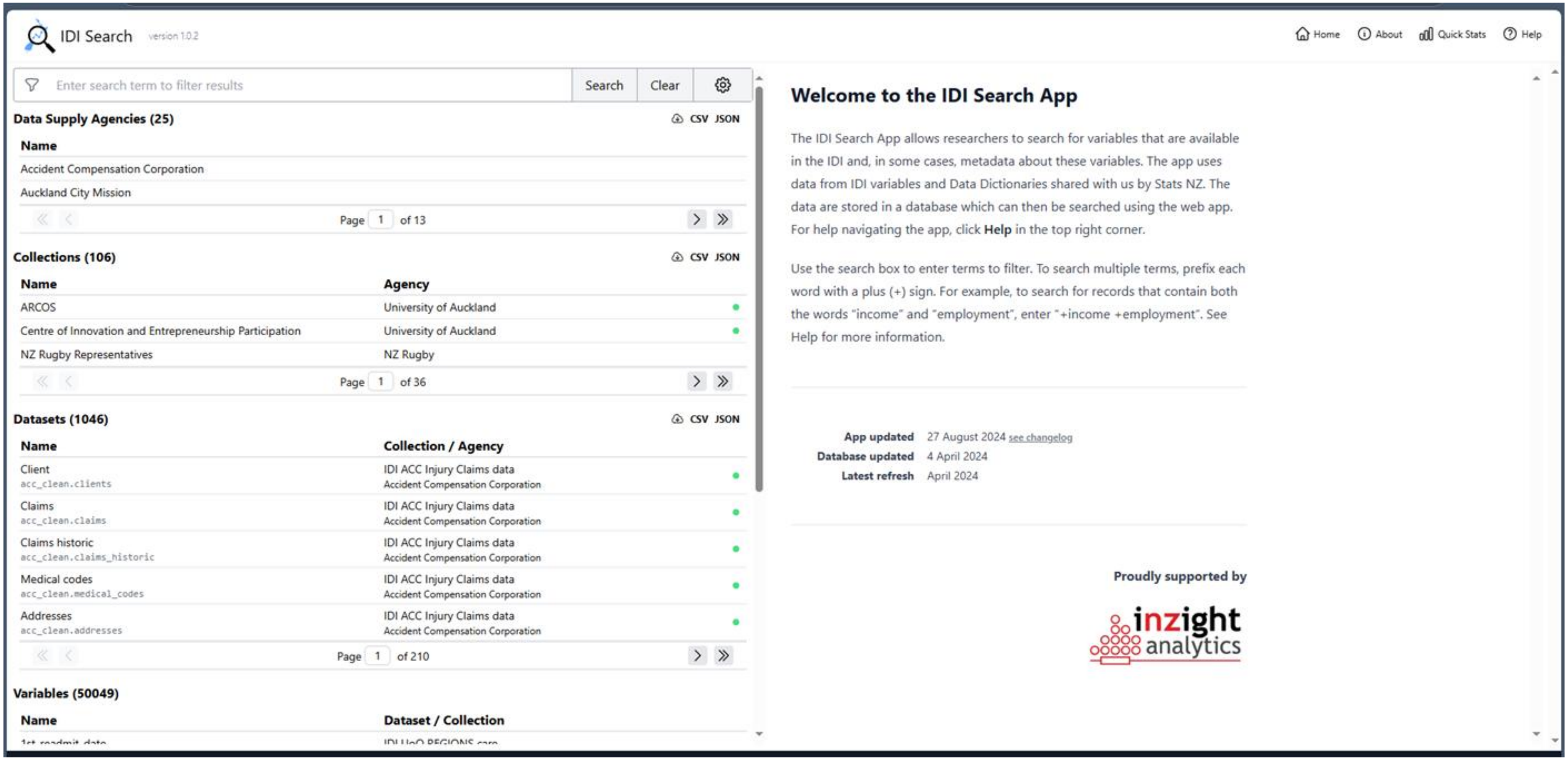
Straightforward as only a single number, benefit receipt, children, and diabetes are all unambiguous concepts.

Analytic approach:

Very challenging.

Diabetes must be constructed from a range of source tables. Multiple ways to define parenting status – may need to test and compare approaches.

Review metadata resources



Data Supply Agencies (25)

Name
Accident Compensation Corporation
Auckland City Mission

Page 1 of 13

Collections (106)

Name	Agency
ARCOS	University of Auckland
Centre of Innovation and Entrepreneurship Participation	University of Auckland
NZ Rugby Representatives	NZ Rugby

Page 1 of 36

Datasets (1046)

Name	Collection / Agency
Client acc_clean.clients	IDI ACC Injury Claims data Accident Compensation Corporation
Claims acc_clean.claims	IDI ACC Injury Claims data Accident Compensation Corporation
Claims historic acc_clean.claims_historic	IDI ACC Injury Claims data Accident Compensation Corporation
Medical codes acc_clean.medical_codes	IDI ACC Injury Claims data Accident Compensation Corporation
Addresses acc_clean.addresses	IDI ACC Injury Claims data Accident Compensation Corporation

Page 1 of 210

Variables (50049)

Name	Dataset / Collection
1st readmit date	IDI UoO REGIONS care

Welcome to the IDI Search App

The IDI Search App allows researchers to search for variables that are available in the IDI and, in some cases, metadata about these variables. The app uses data from IDI variables and Data Dictionaries shared with us by Stats NZ. The data are stored in a database which can then be searched using the web app. For help navigating the app, click **Help** in the top right corner.

Use the search box to enter terms to filter. To search multiple terms, prefix each word with a plus (+) sign. For example, to search for records that contain both the words "income" and "employment", enter "+income +employment". See Help for more information.

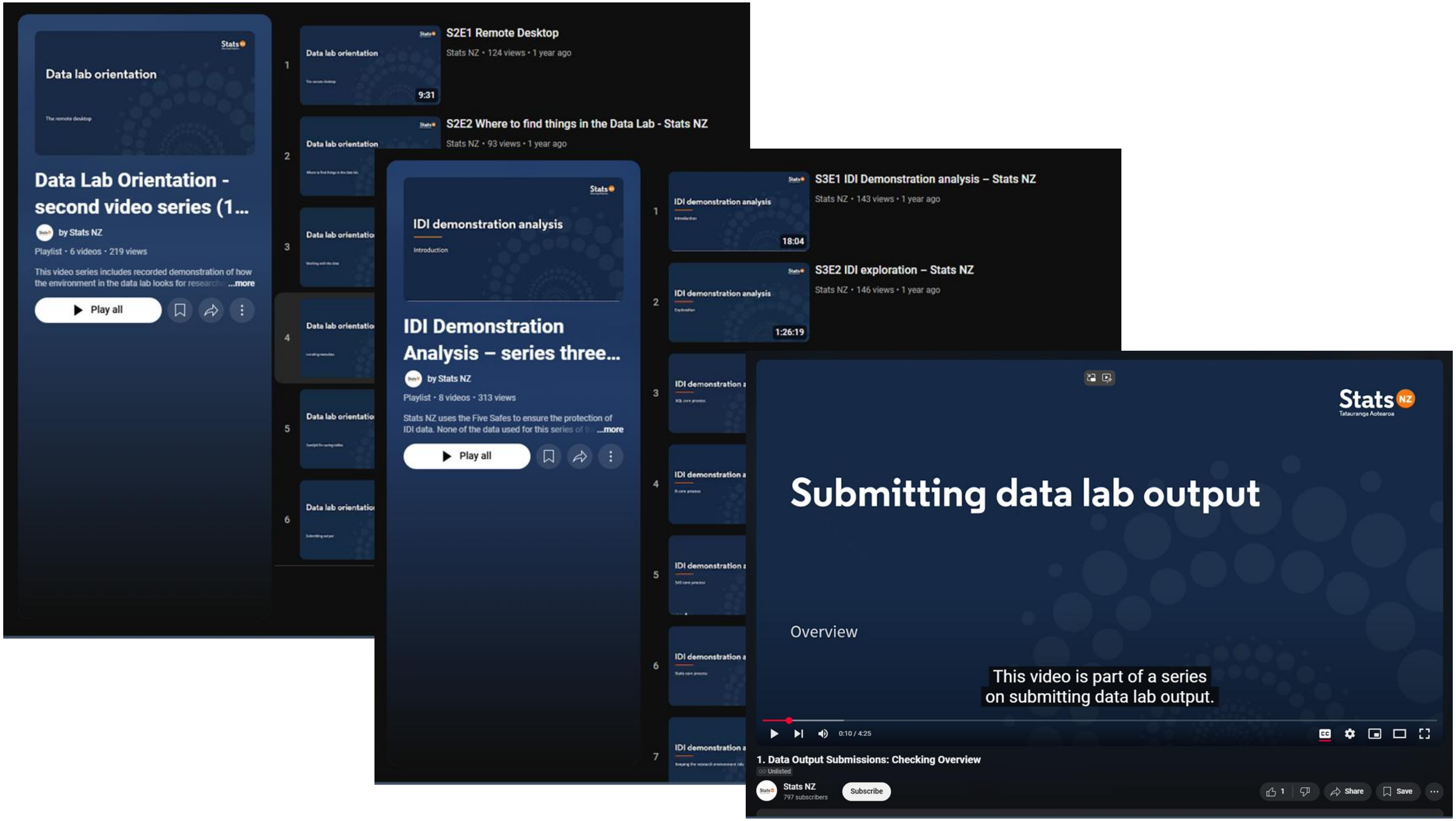
App updated	27 August 2024 see changelog
Database updated	4 April 2024
Latest refresh	April 2024

Proudly supported by

inzight analytics

https://idisearch.terourou.org/

Video series are available



Data Lab Orientation - second video series (1...

by Stats NZ

Playlist • 6 videos • 219 views

This video series includes recorded demonstration of how the environment in the data lab looks for researche ...more

1. S2E1 Remote Desktop
   Stats NZ • 124 views • 1 year ago
2. S2E2 Where to find things in the Data Lab - Stats NZ
   Stats NZ • 93 views • 1 year ago

IDI Demonstration Analysis – series three...

by Stats NZ

Playlist • 8 videos • 313 views

Stats NZ uses the Five Safes to ensure the protection of IDI data. None of the data used for this series of t ...more

1. S3E1 IDI Demonstration analysis – Stats NZ
   Stats NZ • 143 views • 1 year ago
2. S3E2 IDI exploration – Stats NZ
   Stats NZ • 146 views • 1 year ago

Submitting data lab output

Overview

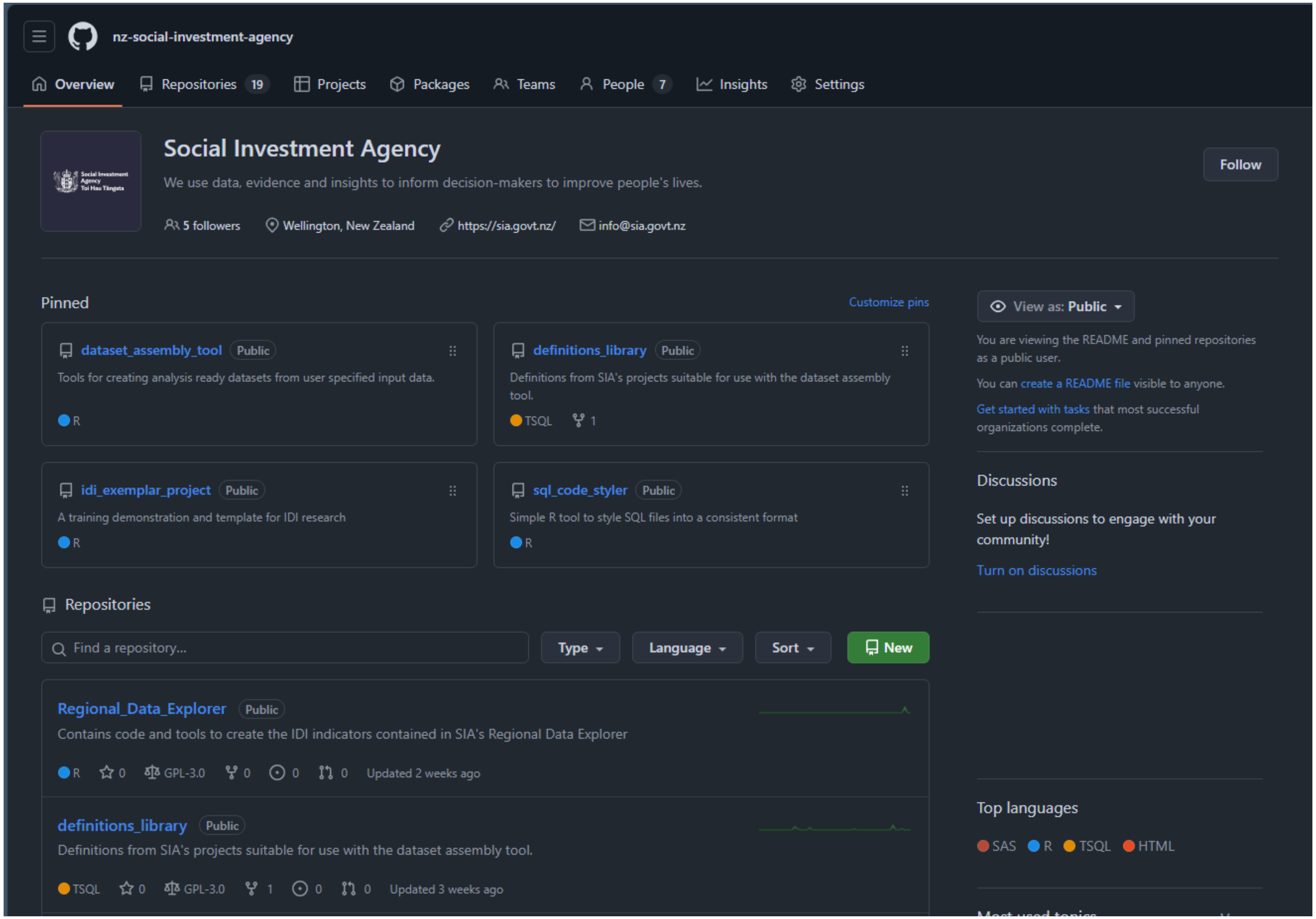
This video is part of a series on submitting data lab output.

1. Data Output Submissions: Checking Overview

Unlisted

Stats NZ
797 subscribers

Build on existing tool and resources



nz-social-investment-agency

Overview | Repositories 19 | Projects | Packages | Teams | People 7 | Insights | Settings

Social Investment Agency

We use data, evidence and insights to inform decision-makers to improve people's lives.

Follow

5 followers · Wellington, New Zealand · https://sia.govt.nz/ · info@sia.govt.nz

Pinned — Customize pins

dataset_assembly_tool Public
Tools for creating analysis ready datasets from user specified input data.
● R

definitions_library Public
Definitions from SIA's projects suitable for use with the dataset assembly tool.
● TSQL 1

idi_exemplar_project Public
A training demonstration and template for IDI research
● R

sql_code_styler Public
Simple R tool to style SQL files into a consistent format
● R

Repositories

Find a repository... | Type | Language | Sort | New

Regional_Data_Explorer Public
Contains code and tools to create the IDI indicators contained in SIA's Regional Data Explorer
● R ☆ 0 GPL-3.0 0 0 0 Updated 2 weeks ago

definitions_library Public
Definitions from SIA's projects suitable for use with the dataset assembly tool.
● TSQL ☆ 0 GPL-3.0 1 0 0 Updated 3 weeks ago

View as: Public

You are viewing the README and pinned repositories as a public user.

You can create a README file visible to anyone.

Get started with tasks that most successful organizations complete.

Discussions

Set up discussions to engage with your community!

Turn on discussions

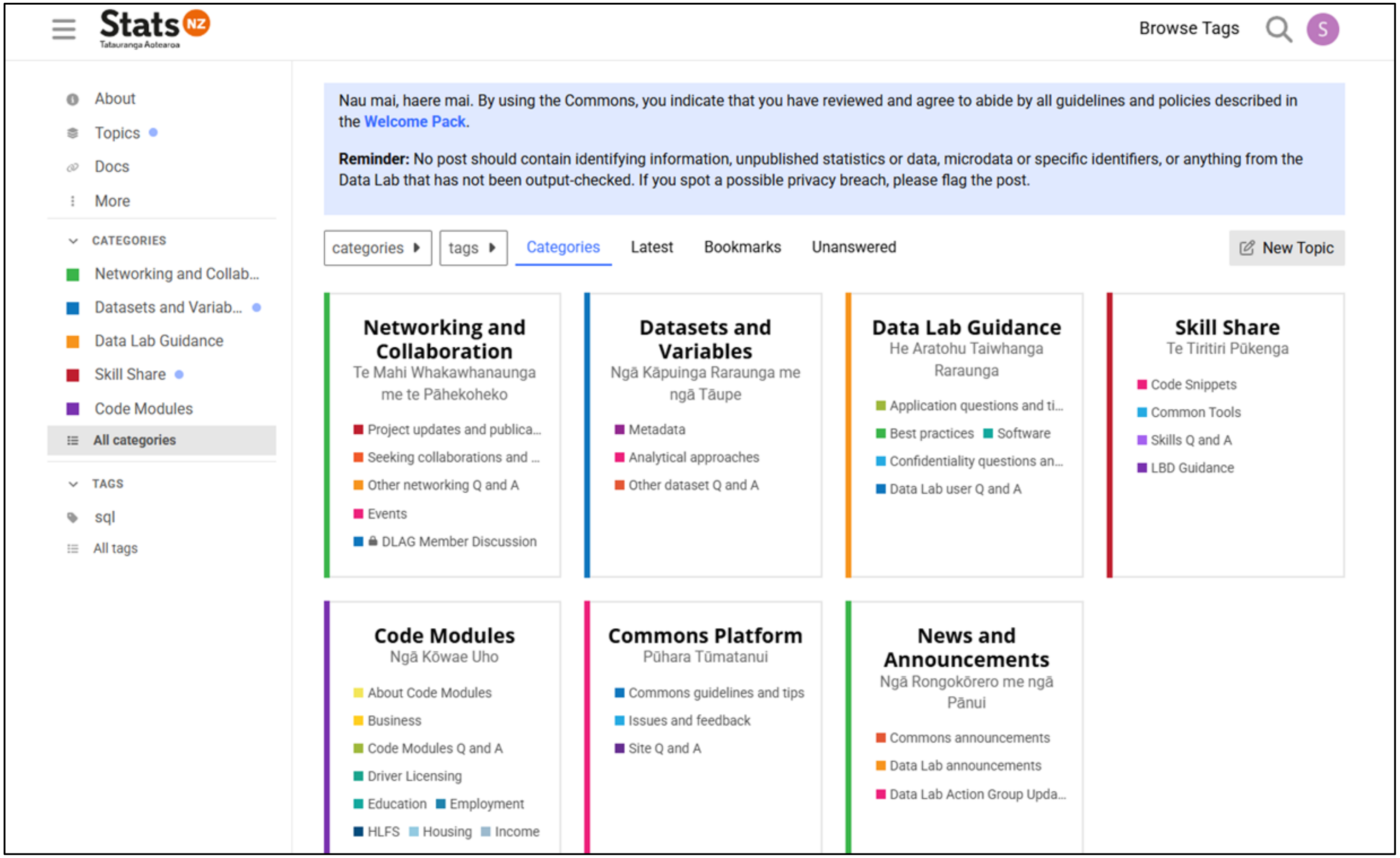
Top languages

● SAS ● R ● TSQL ● HTML

Most used topics

https://github.com/nz-social-investment-agency

Connect with the research community



Stats NZ Tatauranga Aotearoa

Browse Tags

- About
- Topics
- Docs
- More

CATEGORIES

- 🟩 Networking and Collab...
- 🟦 Datasets and Variab...
- 🟧 Data Lab Guidance
- 🟥 Skill Share
- 🟪 Code Modules
- All categories

TAGS

- sql
- All tags

Nau mai, haere mai. By using the Commons, you indicate that you have reviewed and agree to abide by all guidelines and policies described in the **Welcome Pack**.

Reminder: No post should contain identifying information, unpublished statistics or data, microdata or specific identifiers, or anything from the Data Lab that has not been output-checked. If you spot a possible privacy breach, please flag the post.

categories ▸ | tags ▸ | Categories | Latest | Bookmarks | Unanswered | New Topic

Networking and Collaboration
Te Mahi Whakawhanaunga me te Pāhekoheko

- Project updates and publica...
- Seeking collaborations and ...
- Other networking Q and A
- Events
- 🔒 DLAG Member Discussion

Datasets and Variables
Ngā Kāpuinga Raraunga me ngā Tāupe

- Metadata
- Analytical approaches
- Other dataset Q and A

Data Lab Guidance
He Aratohu Taiwhanga Raraunga

- Application questions and ti...
- Best practices
- Software
- Confidentiality questions an...
- Data Lab user Q and A

Skill Share
Te Tiritiri Pūkenga

- Code Snippets
- Common Tools
- Skills Q and A
- LBD Guidance

Code Modules
Ngā Kōwae Uho

- About Code Modules
- Business
- Code Modules Q and A
- Driver Licensing
- Education
- Employment
- HLFS
- Housing
- Income

Commons Platform
Pūhara Tūmatanui

- Commons guidelines and tips
- Issues and feedback
- Site Q and A

News and Announcements
Ngā Rongokōrero me ngā Pānui

- Commons announcements
- Data Lab announcements
- Data Lab Action Group Upda...

https://idcommons.discourse.group/

Plan for confidentiality rules

Design research within rules

- Estimate population size
- List subgroups you want to analyse
- Will every subgroup be large enough?

- Random rounding adds noise
- Will your results be robust with this noise?

- Track entity counts through analysis, difficult to add retrospectively

Make output process easy

- Stats NZ check 50+ submissions every week
- A small amount of extra effort on each output submission adds up to days of extra effort

- Spend a little more time ensuring submission is correct and clear
- Save the checker time and save yourself delays

- Watch the video series on good output practice before your first submission

Distinguish between different table layouts

Tidy rectangular source

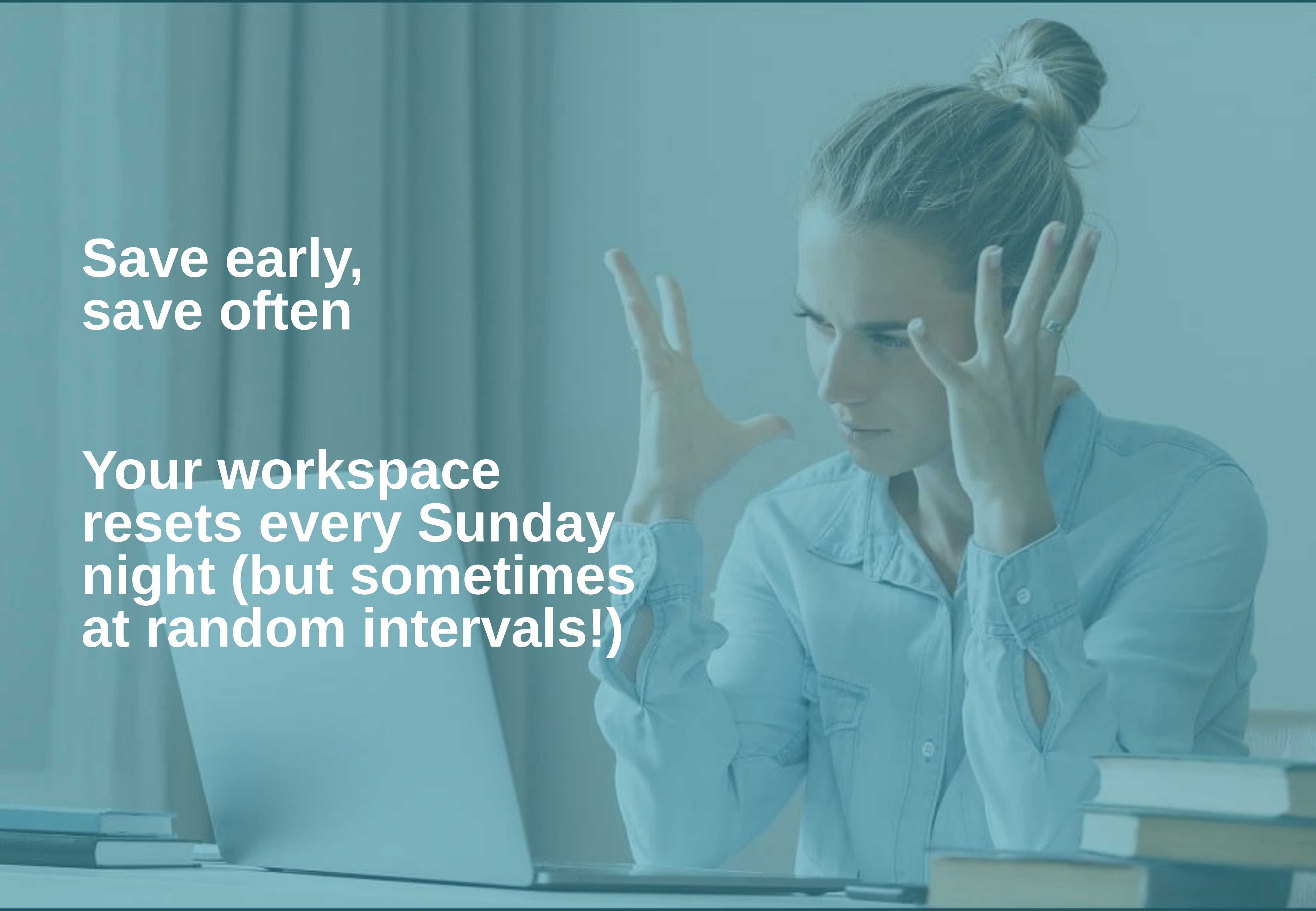
ID	region	age	income
1	north	younger	200
2	north	older	400
3	north	younger	100
4	south	older	200
5	south	younger	100
6	south	older	0
7	south	older	400
8	south	younger	300
9	south	younger	400
10	south	younger	400
11	north	older	100
12	north	older	300
13	north	older	0
14	north	younger	400
15	north	younger	200
16	north	older	300

Long-thin results

region	age	count	total income
north	older	5	1100
north	younger	4	900
south	older	3	600
south	younger	4	1200
-	older	8	1700
-	younger	8	2100
north	-	9	2000
south	-	7	1800
-	-	16	3800

Presentation results

count	younger	older
north	4	5
south	4	3
total income	younger	older
north	900	1100
south	1200	600



Save early,
save often

Your workspace resets every Sunday night (but sometimes at random intervals!)

Simon Anastasiadis

info@sia.govt.nz

sia.govt.nz

Social Investment Agency
Toi Hau Tāngata

Te Kāwanatanga o Aotearoa
New Zealand Government