Open Data, Open Potential

Open Source Code Release

Doug Lambert & Ernestynne Walsh

SIA SOCIAL INVESTMENT AGENCY ORANGA TANGATA

The Social Investment Agency

Investing in what works for better lives

We are a catalyst for change. We apply our skills to technology and data, to create evidence about what works for whom, to improve decision-making and create positive change.

Things we are doing in the Open Data Space

Data Protection & Use Policy

This policy will inform, clarify and guide the use of data for social service delivery, including its use for social investment purposes.

Data Exchange (DX)

A "smart pipe" that enables the safe, secure and easy sharing of data, with a view to improving service effectiveness and outcomes across the social sector. Organisations using the DX retain full control of the what, when and who of their data sharing.

Our Values

Manawa Māui
Catalyst for Change

Taunakitanga
Influence through evidence

He Tangata
About people

Puaretanga
Transparent by nature

Our first open source project

social_investment_analytical_layer

Standardised tables for doing cross agency work in the IDI

social | layer | unit | investment | idi | analytical | siu

SAS ★ 5 4 GPL-3.0 Updated on Jan 6

- Reusable code to save time and money
- Agency recommended derivation of variables
- Code contributions from other agencies
- Available to everyone on GitHub

Case studies

Homes.co.nz

HERE Maps

NZGOAL Software Extension

Thundermaps

Dumpark

Open source software release saves nearly $1m

Continuing our open source journey

social_investment_data_foundation

Framework and modular code for creating an analysis ready dataset in the IDI

SAS 2 GPL-3.0 Updated on Jan 5

social_housing

Repository for the code required to run the Social Housing analysis end-to-end.

SAS 2 GPL-3.0 Updated on Jan 6

mha_data_definition

Standardised definition of mental health and addictions (MHA) service access based on available data in the IDI

SAS 1 GPL-3.0 Updated on Jan 3

sia-vulnerable-mums-analysis

Analysis of natural clusters in a cohort of new mothers, and the isolation of characteristics that contribute to risk of undesirable outcomes.

SAS 1 GPL-3.0 Updated on Jan 4

SIA SOCIAL INVESTMENT AGENCY ORANGA TANGATA

Including open source analytical processes

- Documentation
  - Standardised Headers
  - Inline comments
- Naming Conventions
- Data types
- Formatting of Code
- Folder Structure
- Version Control
- Programming practices
  - Tasks prior to development
  - Macros and functions
  - Improving efficiency
  - Error handling
  - Testing
  - Temporary files
- Graphics

SIA Coding Style Guide

Documentation

Standardised Headers

All headers in scripts must use the standardised format. Examples of populated headers are shown below.

R header

```
# ==============================================================================================
#
# Description: produce plots that compare the K10 pscyhological distress
# measure to the SF-36 perceieved disability measure
#
# Input: [IDI_Sandpit].[DL-MAA2016-15].[k10_vs_mh_band]
#
# Output: comparison_plot = large gg that produces a density plot
# of those with need and those with little or no need
# by average SF36 MH score
#
# Author: E Walsh
#
# Dependencies: k10_sense_check.sas builds the input table that is queried
# this can be run through main_sofie.sas
```

So start your open source journey now

SIA | SOCIAL INVESTMENT AGENCY ORANGA TANGATA