Visualisation driven analytics

Novel ‘point and click’ tools for general and population statistical analysis

Dr. Tom Elliott [1,2] **Daniel Barnett** [2] Andrew Sporle [2,3] Prof. Chris Wild [2]

[1]School of Health, Victoria University of Wellington

[2]Department of Statistics, University of Auckland

[3]iNZight Analytics

‘Point and click’ tools

- lower/remove barrier(s) for entry

‘Point and click’ tools

- lower/remove barrier(s) for entry
- students/researchers/organisations with a limiting resource (time/money/skill)

‘Point and click’ tools

- lower/remove barrier(s) for entry
- students/researchers/organisations with a limiting resource (time/money/skill)
- familiar (drop-downs/sliders/drag-and-drop)

‘Point and click’ tools

- lower/remove barrier(s) for entry
- students/researchers/organisations with a limiting resource (time/money/skill)
- familiar (drop-downs/sliders/drag-and-drop)
- little to no reliance on memory, easy to learn and re-learn

Visualisation driven

- humans are really good at pattern recognition

Visualisation driven

- humans are really good at pattern recognition
- GUI to explore data and look for patterns

Visualisation driven

- humans are really good at pattern recognition
- GUI to explore data and look for patterns
- numerical/textual output/inference comes later

iNZight

What is iNZight?

- a GUI written in R (gWidgets2 + RGtk2)

What is iNZight?

- a GUI written in R (gWidgets2 + RGtk2)
- visualisation-focused tool for exploring/analysing data

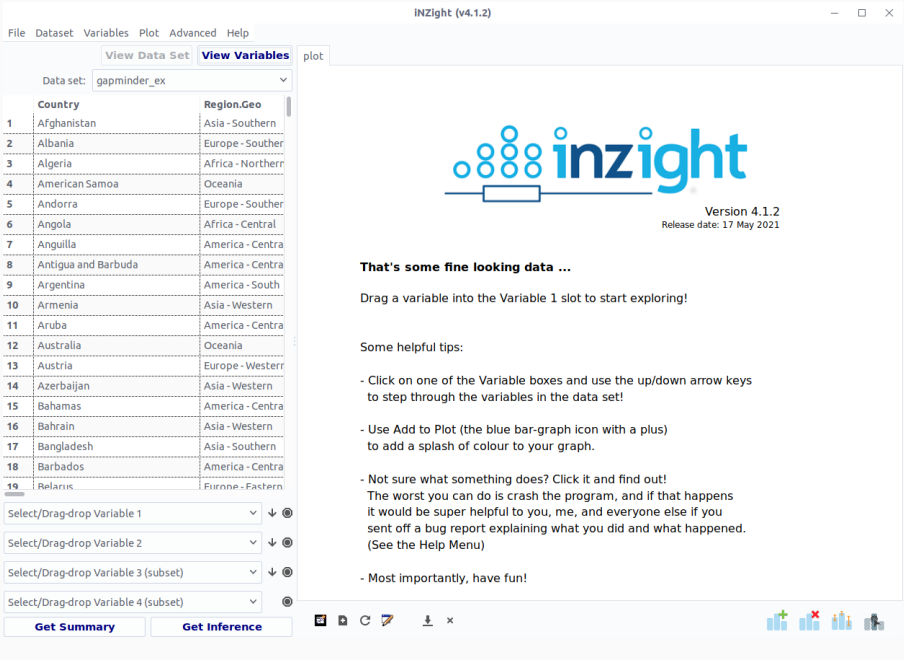
What is iNZight?

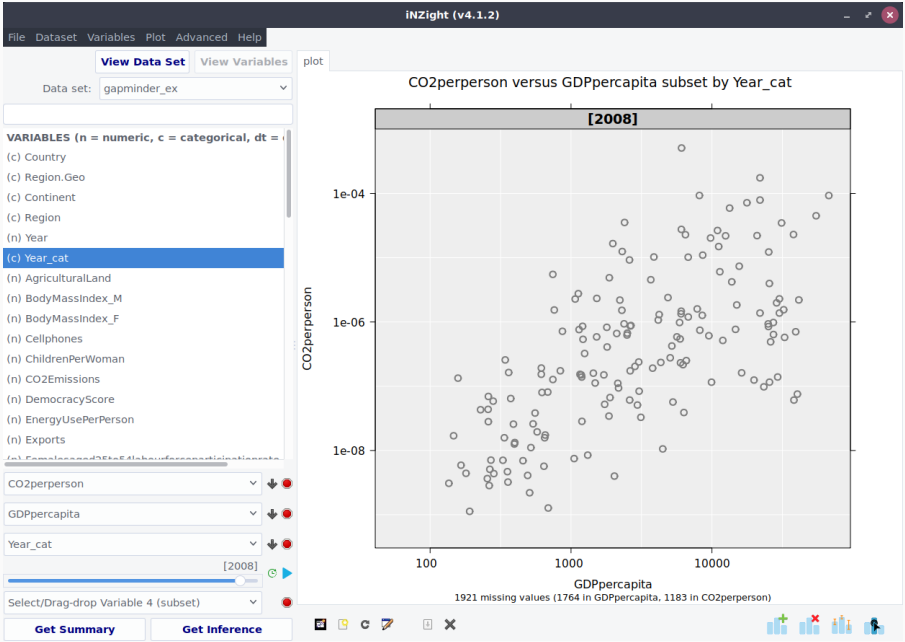
- a GUI written in R (gWidgets2 + RGtk2)
- visualisation-focused tool for exploring/analysing data
- for stats students, orgs/groups with limited resources/time/money

What is iNZight?

- a GUI written in R (gWidgets2 + RGtk2)
- visualisation-focused tool for exploring/analysing data
- for stats students, orgs/groups with limited resources/time/money
- extensions: time series, maps, regression modelling, multiple response, multivariate, Bayesian demography ...

iNZight (v4.1.2)
File Dataset Variables Plot Advanced Help
View Data Set
View Variables
plot
Data set:
data
empty
1
inzight
Version 4.1.2
Release date: 17 May 2021
Kia ora and welcome! To get started, import some data.
Not sure where to go? Try the File menu!
There are some example datasets there if you just want to explore the program.
We've gone ahead and made a few improvements ...
- Code writing for plots, summaries, and inference!
It's not enabled by default, but the biggest change in this release
is the implementation of interactive code snippets, enabled from
File > Preferences. You can see the R code for the current plot,
and edit it to see what happens. Learning R just got easier!
- Get Summary and Get Inference have been redesigned
Now more consistent with the rest of iNZight, you can quickly add and
remove inference information, adjust your hypothesis, and see the
results instantly.
- Survey designs are now completely supported: plots, summaries,
and data wrangling!
- Plus other bug fixes and improvements.
Select/Drag-drop Variable 1
Select/Drag-drop Variable 2
Select/Drag-drop Variable 3 (subset)
Select/Drag-drop Variable 4 (subset)
Get Summary
Get Inference


iNZight (v4.1.2)
File Dataset Variables Plot Advanced Help
View Data Set
View Variables
plot
Data set: gapminder_ex
Country
Region.Geo
1 Afghanistan Asia - Southern
2 Albania Europe - Souther
3 Algeria Africa - Northern
4 American Samoa Oceania
5 Andorra Europe - Souther
6 Angola Africa - Central
7 Anguilla America - Centra
8 Antigua and Barbuda America - Centra
9 Argentina America - South
10 Armenia Asia - Western
11 Aruba America - Centra
12 Australia Oceania
13 Austria Europe - Western
14 Azerbaijan Asia - Western
15 Bahamas America - Centra
16 Bahrain Asia - Western
17 Bangladesh Asia - Southern
18 Barbados America - Centra
19 Belarus Europe - Eastern
Select/Drag-drop Variable 1
Select/Drag-drop Variable 2
Select/Drag-drop Variable 3 (subset)
Select/Drag-drop Variable 4 (subset)
Get Summary
Get Inference
inzight
Version 4.1.2
Release date: 17 May 2021
That's some fine looking data ...
Drag a variable into the Variable 1 slot to start exploring!
Some helpful tips:
- Click on one of the Variable boxes and use the up/down arrow keys to step through the variables in the data set!
- Use Add to Plot (the blue bar-graph icon with a plus) to add a splash of colour to your graph.
- Not sure what something does? Click it and find out! The worst you can do is crash the program, and if that happens it would be super helpful to you, me, and everyone else if you sent off a bug report explaining what you did and what happened. (See the Help Menu)
- Most importantly, have fun!


iNZight (v4.1.2)
File Dataset Variables Plot Advanced Help
View Data Set
View Variables
plot
Data set:
gapminder_ex
VARIABLES (n = numeric, c = categorical, dt =
(c) Country
(c) Region.Geo
(c) Continent
(c) Region
(n) Year
(c) Year_cat
(n) AgriculturalLand
(n) BodyMassIndex_M
(n) BodyMassIndex_F
(n) Cellphones
(n) ChildrenPerWoman
(n) CO2Emissions
(n) DemocracyScore
(n) EnergyUsePerPerson
(n) Exports
CO2perperson
GDPpercapita
Year_cat
[2008]
Select/Drag-drop Variable 4 (subset)
Get Summary
Get Inference
CO2perperson versus GDPpercapita subset by Year_cat
[2008]
1e-04
1e-06
1e-08
CO2perperson
100
1000
10000
GDPpercapita
1921 missing values (1764 in GDPpercapita, 1183 in CO2perperson)

iNZight (v4.1.2)

File Dataset Variables Plot Advanced Help

Add to Plot :

Customise Plot Appearance

General Appearance

Plot type : scatter

Background colour : lightgrey

Overall size scale : 0.8

Point Size

Overall : 1.6

Resize points by : Populationdensity

Resize method : proportional

Points area proportional to value of variable.

Point Colour

Point colour : contrast (max 8)

Colour by : Region

Reverse palette

Cycle levels :

Transparency : 61

Point Symbol

Match with colour variable

(requires categorical variable with 5 or fewer levels)

Symbol : circle

Symbol line width : 2

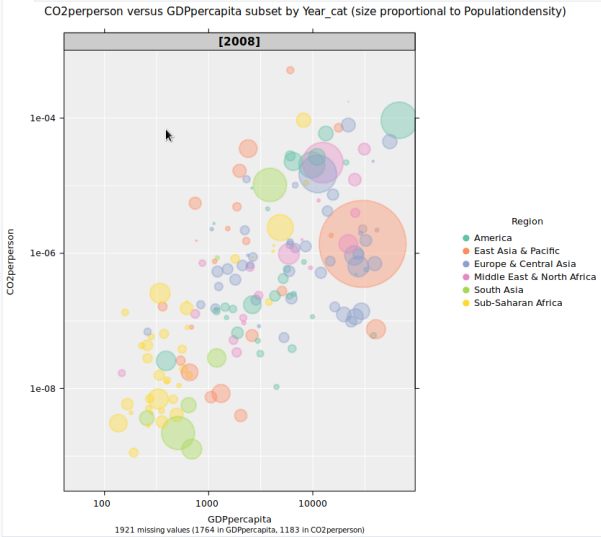
Fill symbols

Update automatically

? Help

Home

plot



Visualisation driven: show *then* tell

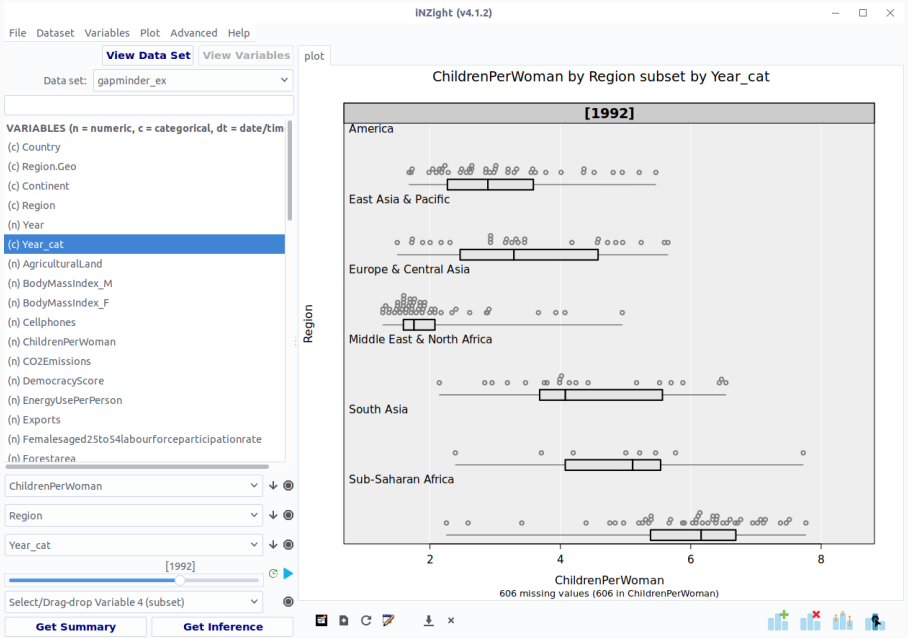
- automatic graphics based on variable types (numeric, categorical, date/time)

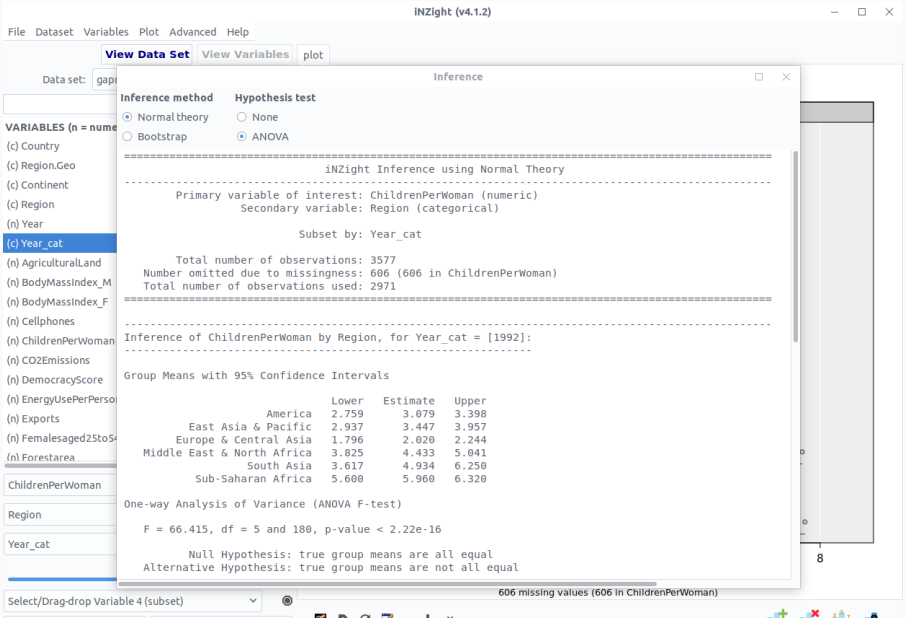
Visualisation driven: show *then* tell

- automatic graphics based on variable types (numeric, categorical, date/time)
- encourages exploration of the data (colour, size, faceting)

Visualisation driven: show *then* tell

- automatic graphics based on variable types (numeric, categorical, date/time)
- encourages exploration of the data (colour, size, faceting)
- summary/inference/hypothesis testing for the current graph


iNZight (v4.1.2)
File Dataset Variables Plot Advanced Help
View Data Set
View Variables
plot
Data set: gapminder_ex
VARIABLES (n = numeric, c = categorical, dt = date/tim
(c) Country
(c) Region.Geo
(c) Continent
(c) Region
(n) Year
(c) Year_cat
(n) AgriculturalLand
(n) BodyMassIndex_M
(n) BodyMassIndex_F
(n) Cellphones
(n) ChildrenPerWoman
(n) CO2Emissions
(n) DemocracyScore
(n) EnergyUsePerPerson
(n) Exports
(n) Femalesaged25to54labourforceparticipationrate
(n) Forestarea
ChildrenPerWoman
Region
Year_cat
[1992]
Select/Drag-drop Variable 4 (subset)
Get Summary
Get Inference
ChildrenPerWoman by Region subset by Year_cat
[1992]
America
East Asia & Pacific
Europe & Central Asia
Middle East & North Africa
South Asia
Sub-Saharan Africa
Region
2
4
6
8
ChildrenPerWoman
606 missing values (606 in ChildrenPerWoman)


iNZight (v4.1.2)
File Dataset Variables Plot Advanced Help
View Data Set
View Variables
plot
Data set:
VARIABLES (n = nume
(c) Country
(c) Region.Geo
(c) Continent
(c) Region
(n) Year
(c) Year_cat
(n) AgriculturalLand
(n) BodyMassIndex_M
(n) BodyMassIndex_F
(n) Cellphones
(n) ChildrenPerWoman
(n) CO2Emissions
(n) DemocracyScore
(n) EnergyUsePerPerso
(n) Exports
(n) Femalesaged25toS4
(n) Forestarea
ChildrenPerWoman
Region
Year_cat
Select/Drag-drop Variable 4 (subset)
Get Summary
Get Inference
Inference
Inference method
Normal theory
Bootstrap
Hypothesis test
None
ANOVA
==========
iNZight Inference using Normal Theory

Primary variable of interest: ChildrenPerWoman (numeric)
Secondary variable: Region (categorical)
Subset by: Year_cat
Total number of observations: 3577
Number omitted due to missingness: 606 (606 in ChildrenPerWoman)
Total number of observations used: 2971
==========

Inference of ChildrenPerWoman by Region, for Year_cat = [1992]:

Group Means with 95% Confidence Intervals
Lower Estimate Upper
America 2.759 3.079 3.398
East Asia & Pacific 2.937 3.447 3.957
Europe & Central Asia 1.796 2.020 2.244
Middle East & North Africa 3.825 4.433 5.041
South Asia 3.617 4.934 6.250
Sub-Saharan Africa 5.600 5.960 6.320
One-way Analysis of Variance (ANOVA F-test)
F = 66.415, df = 5 and 180, p-value < 2.22e-16
Null Hypothesis: true group means are all equal
Alternative Hypothesis: true group means are not all equal
606 missing values (606 in ChildrenPerWoman)
8

Lowering the barrier to entry

- free

Lowering the barrier to entry

- free
- familiar/intuitive controls (drop-downs, sliders)

Lowering the barrier to entry

- free
- familiar/intuitive controls (drop-downs, sliders)
- advanced techniques (faceting, trend lines, hypothesis testing, etc)

Lowering the barrier to entry

- free
- familiar/intuitive controls (drop-downs, sliders)
- advanced techniques (faceting, trend lines, hypothesis testing, etc)
- complex survey designs handled natively

Lowering the barrier to entry

- free
- familiar/intuitive controls (drop-downs, sliders)
- advanced techniques (faceting, trend lines, hypothesis testing, etc)
- complex survey designs handled natively
- R code writing for learning, reproducibility, and research development

inzight.nz

SPONSORS

University of Auckland, Statistics NZ, Ministry of Education, Australian Bureau of Statistics, Te Rourou Tātaritanga (MBIE-funded research), iNZight Analytics

Mātau

What is Mātau?

- a point-and-click app designed for outcome visualisation

What is Mātau?

- a point-and-click app designed for outcome visualisation
- automation of tedious tasks (orgs/countries with limited resources)

What is Mātau?

- a point-and-click app designed for outcome visualisation
- automation of tedious tasks (orgs/countries with limited resources)
- default options are those typically used in official reports

What is Mātau?

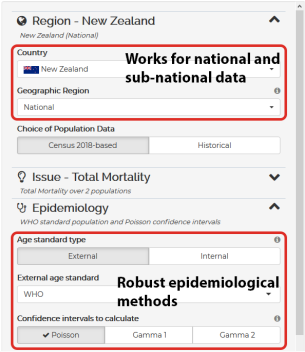
- a point-and-click app designed for outcome visualisation
- automation of tedious tasks (orgs/countries with limited resources)
- default options are those typically used in official reports
- expert options are displayed in pull-down menus

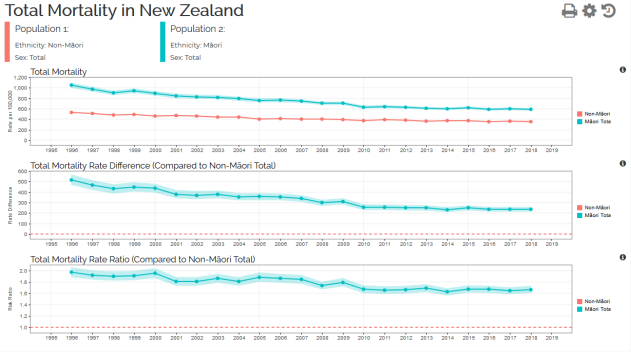
What is Mātau?

- a point-and-click app designed for outcome visualisation
- automation of tedious tasks (orgs/countries with limited resources)
- default options are those typically used in official reports
- expert options are displayed in pull-down menus
- interactive plots using D3

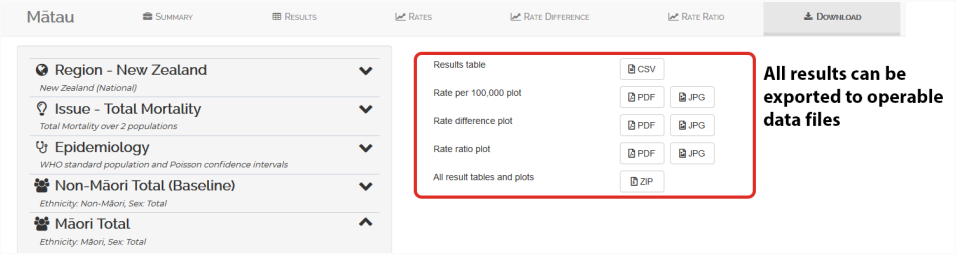
A first view of Mātau

Options hidden behind pull-down menus

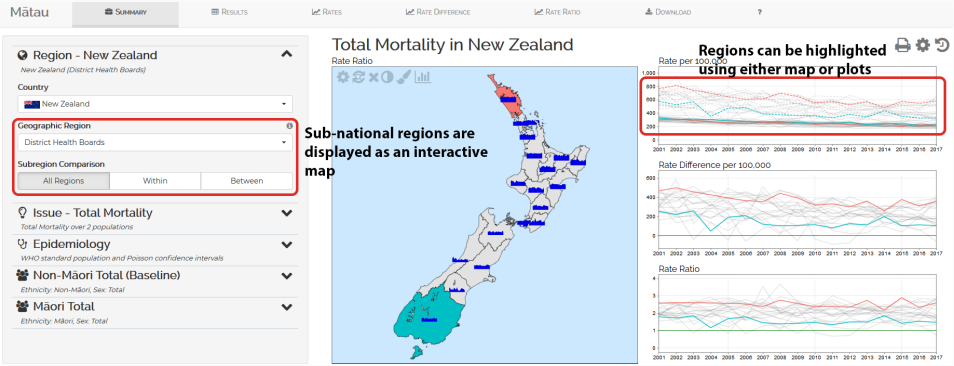




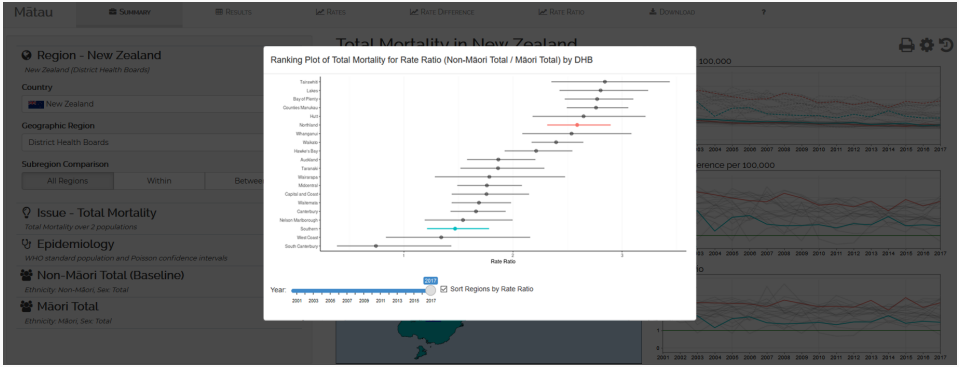
Exportable results



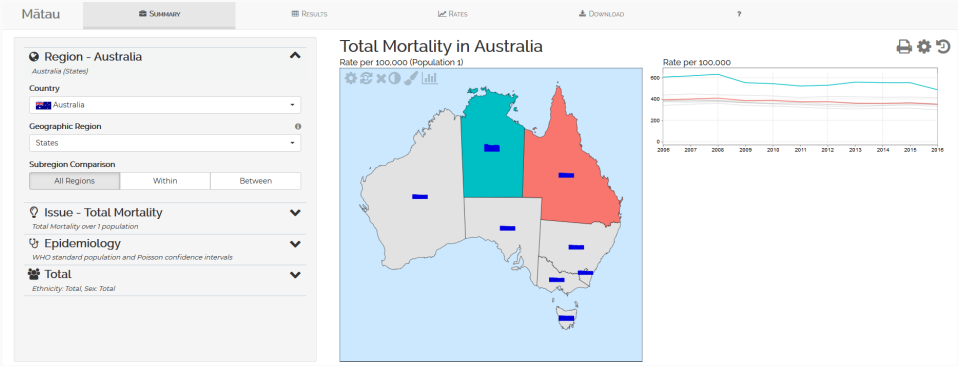
Sub-national estimates



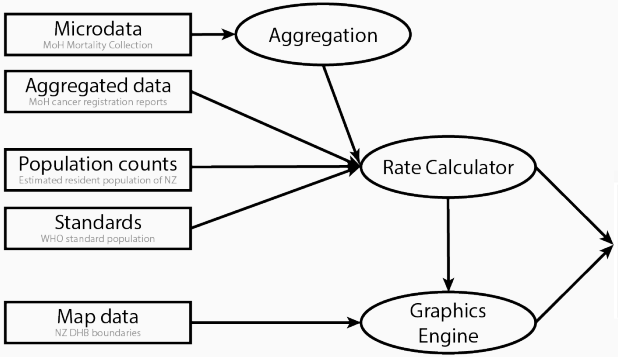
Sub-national estimates



Mātau even works with Australian data



Behind the scenes of Mātau



Tom Elliott

Victoria University of Wellington
University of Auckland

tom.elliott@{vuw,auckland}.ac.nz

Daniel Barnett

University of Auckland

daniel.barnett@auckland.ac.nz

inzight

https://inzight.nz

Mātau

Website to come soon

THE UNIVERSITY OF AUCKLAND NEW ZEALAND

VICTORIA UNIVERSITY OF WELLINGTON TE HERENGA WAKA

TE ROUROU TĀTARITANGA
Informatics for Social Services and Wellbeing

MINISTRY OF BUSINESS, INNOVATION & EMPLOYMENT
HIKINA WHAKATUTUKI

inzight analytics