New Zealand's journey toward an admin-first 'census' future

UN-ECE – 2023 Census Week – 18-22 Sep 2023

Rochelle Morgan, Senior Manager – Population Measurement
& 2023 Census Methodology Lead, Methods & Design, Stats NZ

TĀTOU

1. (verb) to count, calculate, add up, enumerate, tally

2. (noun) number, count, figure, tally, total, score

Outline

- Our context – New Zealand census & history of census transformation
- Our research programme – our admin data & where quality is at
- Our combined census – the 2023 Census 'by design'
- Our environment – legislation & Māori expectations
- Our transformation – the admin-first future is here…
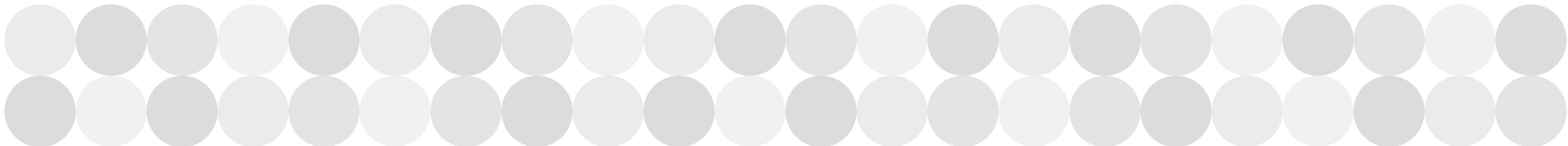
Our context…

The New Zealand five yearly census

Stats NZ
Tatauranga Aotearoa

2011: Drivers for change

Increasing cost of Census

Sustainability of current model

Opportunities from new technologies

Availability of admin data

Govt. push for integrated data to improve govt. services

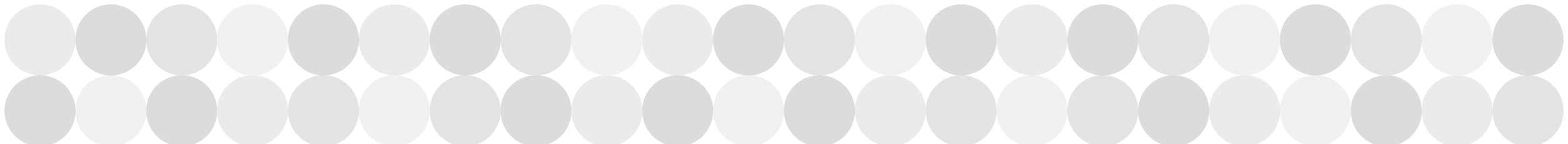*2012: Census Transformation Strategy (our research programme):*

Vision: *Administrative census, with survey support*

Short-medium term: modernise current census model → Long term: investigate census based on administrative data

New Zealand Census & use of administrative data

Stats NZ
Tatauranga Aotearoa

Research programme

Year	Event
2011	2011 Census cancelled due to Christchurch earthquake
2012	Census transformation strategy agreed with government
2013	2013 Census (delayed) – traditional full field enumeration model
2015	Rapid expansion of Stats NZ's Integrated Data Infrastructure	Stats NZ (IDI)
2018	2018 Census – planned modernised traditional model, combined model as mitigation
2023	2023 Census – respondent-first combined model 'by design'
2028	The next census will be admin-first…
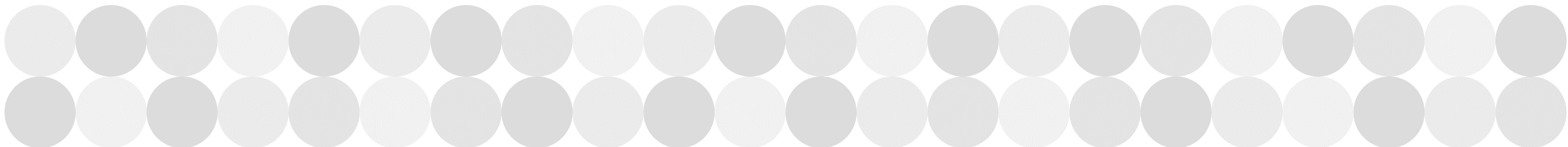
Our research programme…

Using admin data means a transformation

Stats NZ
Tatauranga Aotearoa

Quality of admin data:

How do you know when admin data is good enough?

| Understand your sources | | Define your statistical |

| Administrative data | What are the strengths & limitations?

What transformations are needed? | Statistical uses |

Use of admin data:
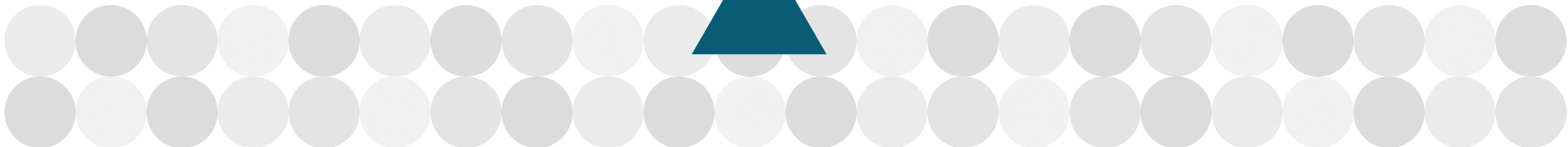
Balancing privacy & benefits

Right to privacy

Right to live in an informed society

Can we?   Should we?

Administration Resident Population

Stats NZ
Tatauranga Aotearoa
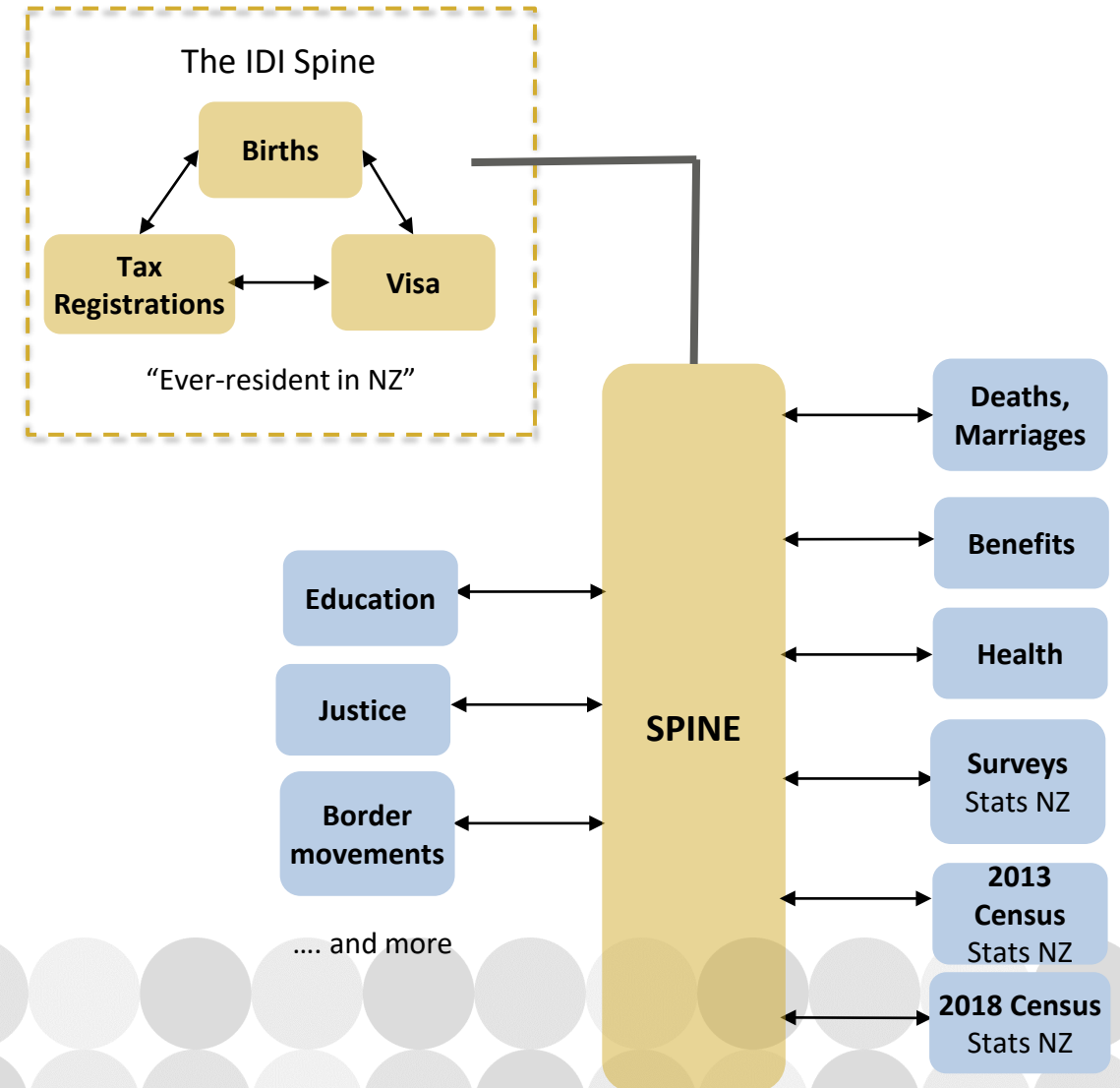
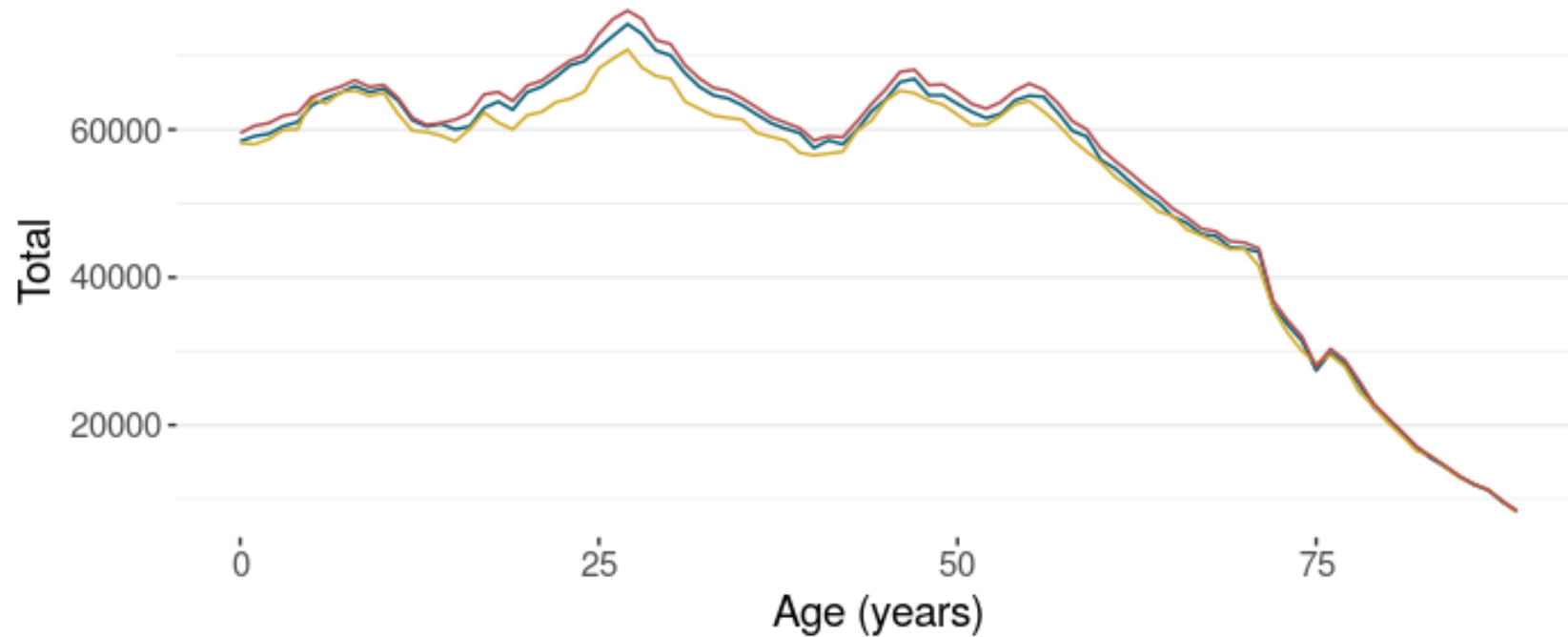Linked administrative data:
Integrated Data Infrastructure (IDI)

Administrative NZ resident population
- **Begin** with IDI spine ('ever-resident' population)
  - **Include** all individuals with **activity** in admin data sources (tax and benefits, health, education) within the previous two years
  - **Remove** individuals
    - who died before reference date
    - who migrated overseas before reference date
    - remove the non-residents (border movements data)

The IDI Spine

Births

Tax Registrations ↔ Visa

"Ever-resident in NZ"

Education

Justice

Border movements

.... and more

SPINE

Deaths, Marriages

Benefits

Health

Surveys Stats NZ

2013 Census Stats NZ

2018 Census Stats NZ

What does the admin resident population look like?

NZ usual resident population,
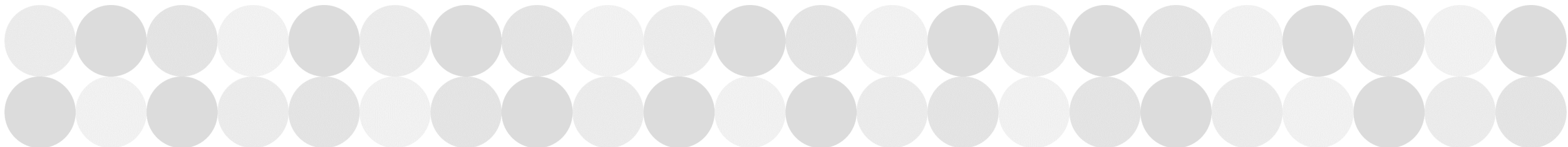for APC and ERP at 30 June 2018 and 2018 Census

Admin resident population (APC) is close to official estimated resident population (ERP).

- Small undercount, fairly even across ages
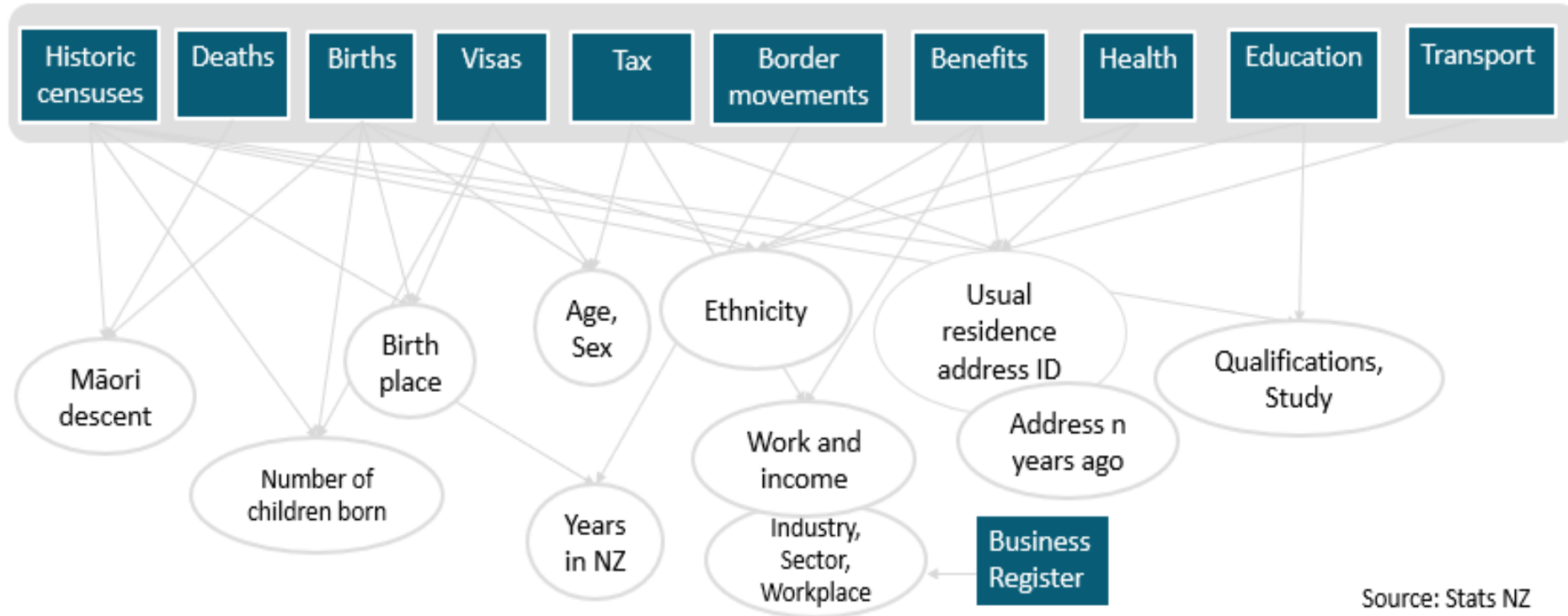- Admin population is higher than the census counts

Source: Stas NZ
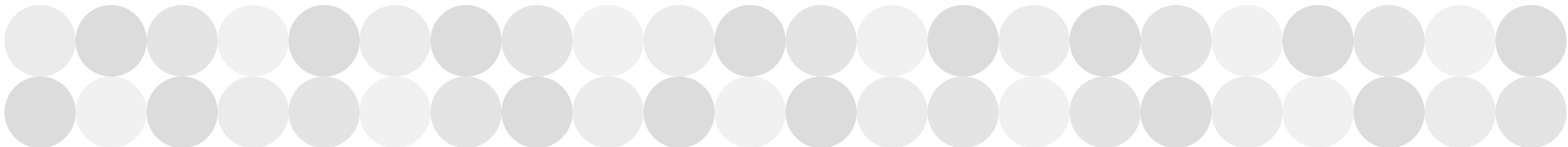
Where the administrative data sources come from

![Stats NZ Tatauranga Aotearoa]

All data sources are linked in and sourced from the Stats NZ Integrated Data Infrastructure (IDI)

IDI Spine:

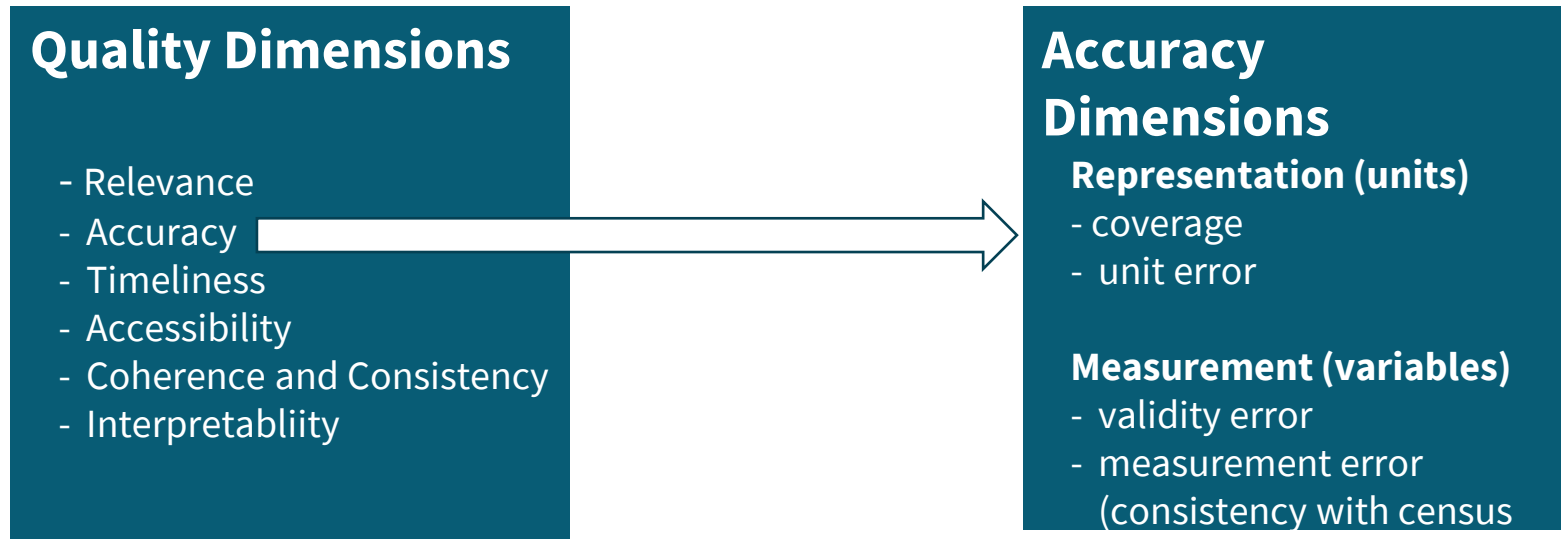| Historic censuses | Deaths | Births | Visas | Tax | Border movements | Benefits | Health | Education | Transport |

Around 10m individuals 'ever resident in NZ'

Māori descent

Birth place

Age, Sex

Ethnicity

Usual residence address ID

Qualifications, Study

Number of children born

Years in NZ

Work and income

Address n years ago

Industry, Sector, Workplace

Business Register

Source: Stats NZ

Admin data quality frameworks guide assessment of variables

Quality Dimensions

- Relevance
- Accuracy
- Timeliness
- Accessibility
- Coherence and Consistency
- Interpretabliity

Accuracy Dimensions

Representation (units)
- coverage
- unit error

Measurement (variables)
- validity error
- measurement error (consistency with census

Statistics New Zealand (2016) Guide to reporting on administrative data quality | Stats NZ Zhang, L-C (2012). "Topics of statistical theory for register-based statistics and data integration." Statistica Neerlandica 66: 41–63
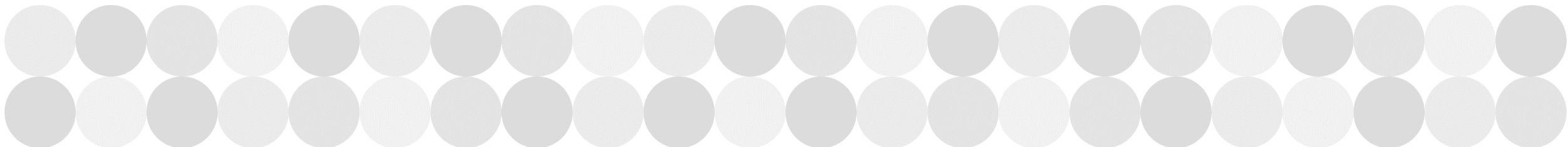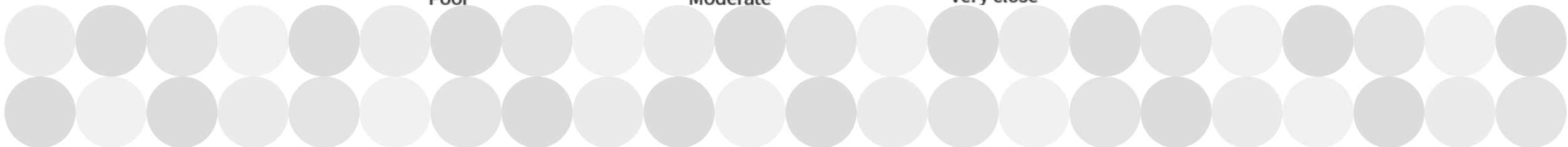
Quality assessment of admin sources
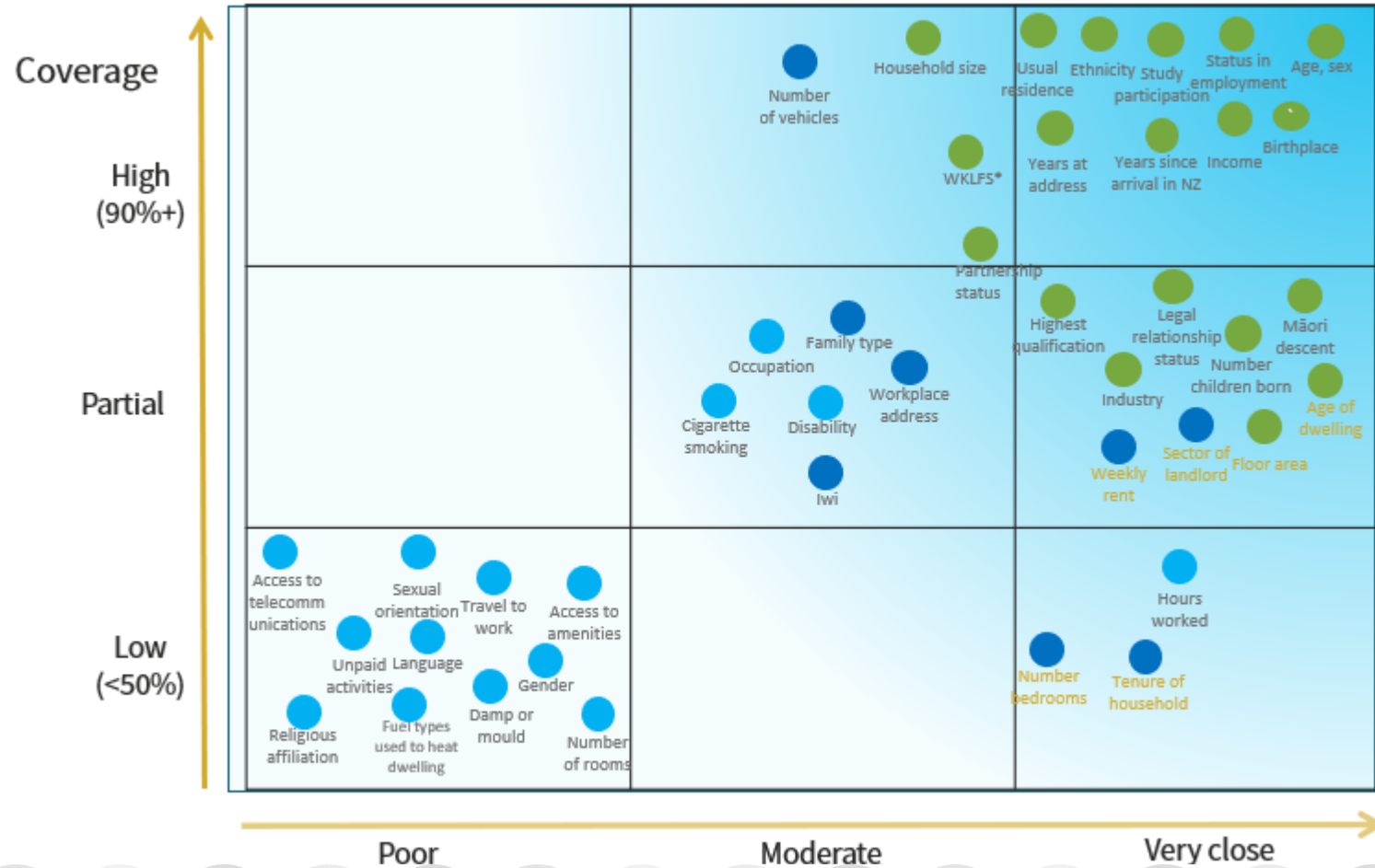
As at Aug 2023
- subject to change

Metadata analysis

Detailed analysis

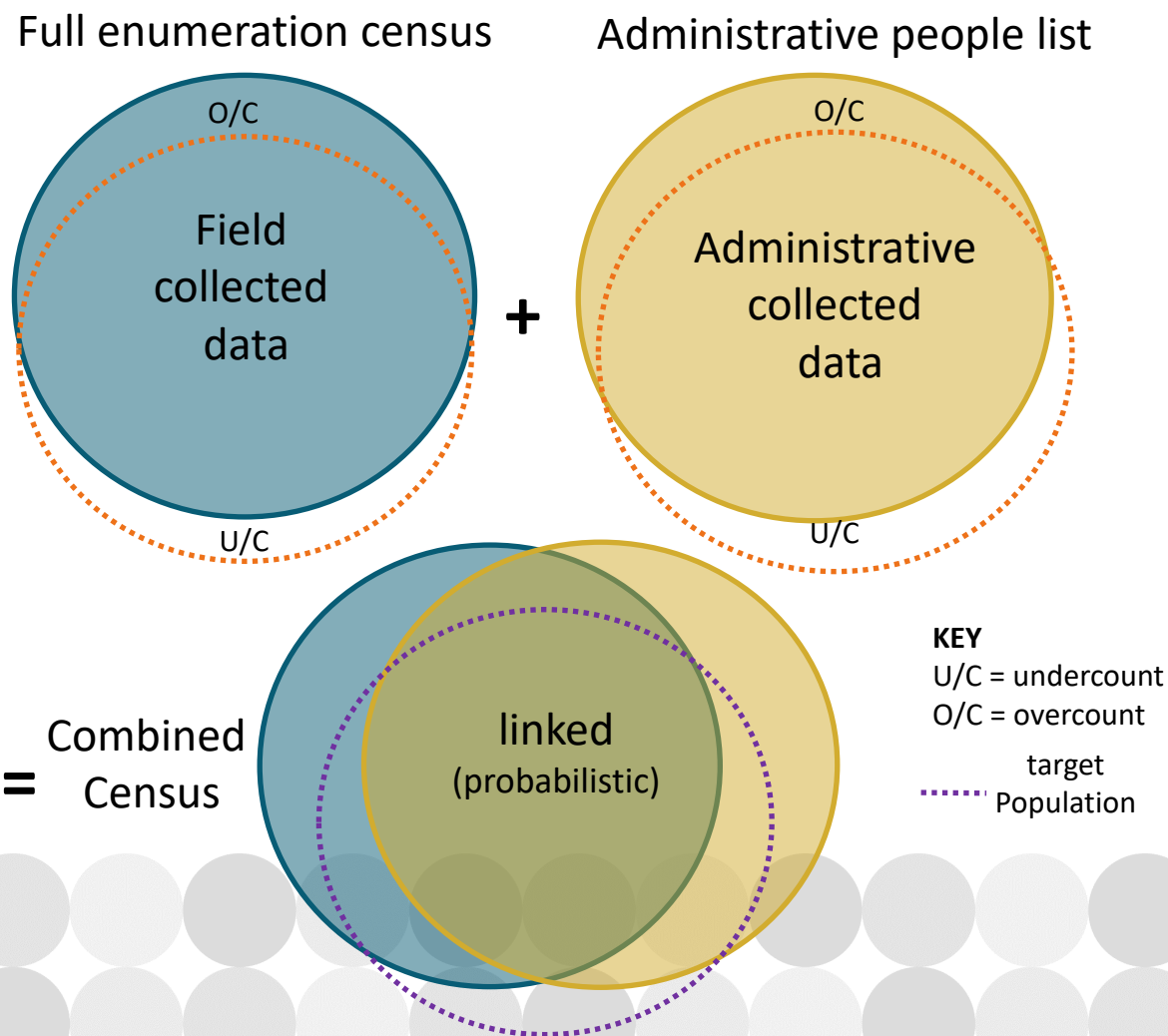Output in experimental *Admin Population Census (APC)*

Our combined census...

The 2023 Combined Census 'by design'

Combined Census 'by design' – admin data scope statements:

"Administrative and alternative data sources will be used to make improvements to existing data processes and output quality, to reduce or resolve known issues in processing, and to support the processing of data under a combined model approach.

The specific uses of administrative and alternative data sources for units and characteristics will be determined by the 'combined model' design.

Administrative and alternative data sources will also be used in the development of interim population estimates to support analysis and quality assurance."
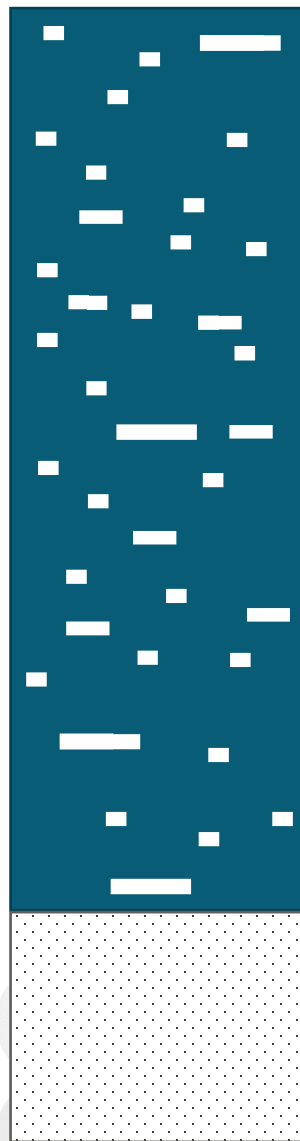

Full enumeration census

O/C

Field collected data

U/C

+

Administrative people list

O/C

Administrative collected data

U/C

= Combined Census

linked (probabilistic)

KEY
U/C = undercount
O/C = overcount

target Population

From census responses to the combined census

Stats NZ
Tatauranga Aotearoa

The combined census model 'by design' combines an admin population list, equivalent to the [Admin Population Census](#) (APC), with census collected response data to achieve high coverage and quality, including rich attributes information.

Collection

2023 Census Responses

Add alternative data
(for units and attributes)

from census responses to the final combined census dataset

Data Delivery

Final census unit record file (CURF)

Admin enumerations

■ Census response
■ Admin data*
■ Statistical Imputation
□ Missing information
▱ Undercount

Use alternative data to fill in missing information

Stats NZ Tatauranga Aotearoa

No previous census or admin data available: Use statistical imputation?

Use previous census and admin data to fill in missing information; Statistical imputation if neither available
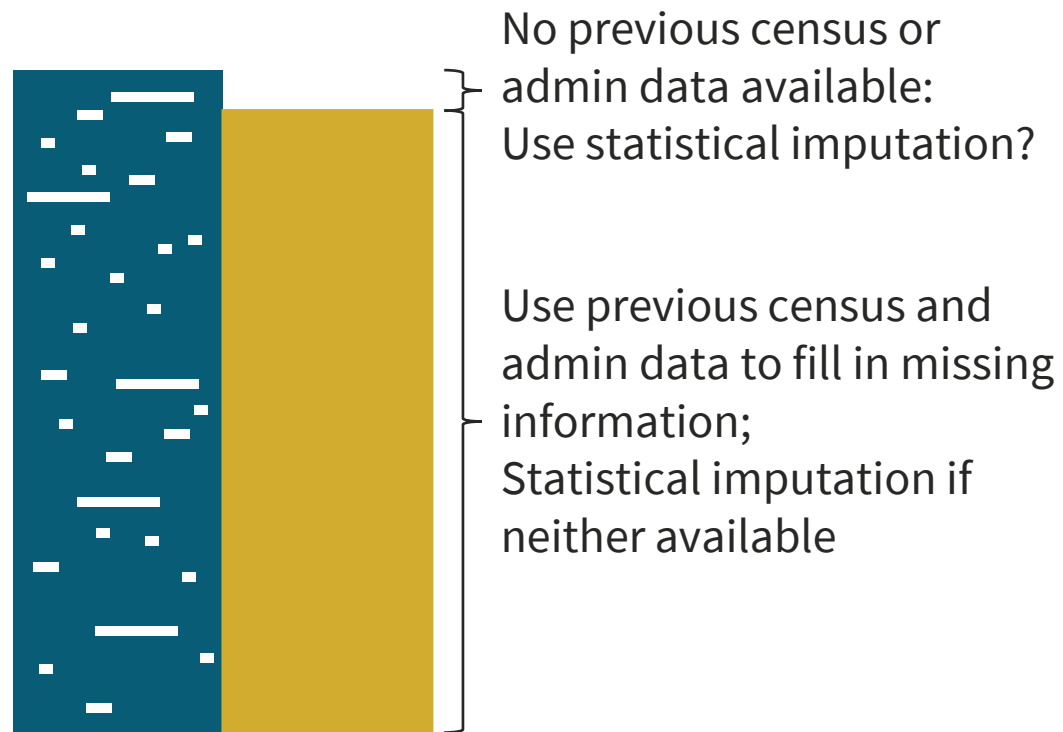
Prioritised sources for filling in a variable:

2023 Census Response

↓

Response from 2013 or 2018 Census?

↓

Admin Data?

↓

Statistical Imputation?

Our future…

Our future:

Stats NZ — Tatauranga Aotearoa

Drivers for change:

- Changing data landscape
- Data collection challenges
- Low response rates
- Māori data sovereignty
- Demands on data
- Expectations on data processing power
- Democratisation of data

Our future:

- Designing with communities and partners
- Measure what matters; trust and confidence
- Increased use of admin data
- Simplify surveys
- New technologies
- Lead government data system

Goal: *Become an organisation that uses administrative data first*

Admin first priorities:

- Ensuring our digital infrastructure sets us up for the future incl. statistical registers
- Researching, designing and developing methodologies to inform future census models
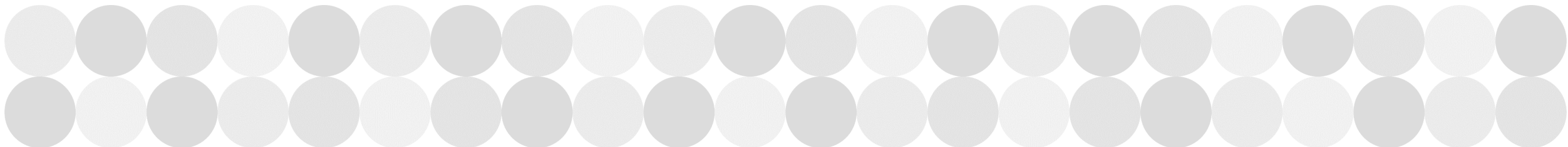- Working with data suppliers, our te Tiriti partners
- Ensure our internal capability and operating model sets us up for the future

Our environment – Data & Statistics Act 2022

Modernised legislation

(previously the Statistics Act 1975)

It's now easier for Stats NZ to make use of public sector data for official statistics:

- The Statistician can <u>require</u> a public sector agency to provide data

- Agencies must proactively <u>inform</u> us of changes to data or concerns about <u>quality</u>

- Requires Census to be five-yearly but is <u>non-prescriptive</u> of how the data is collected

Our new Act speaks directly to the requirement for the Crown to give effect to the principles of te Tiriti o Waitangi / the Treaty of Waitangi

Data and Statistics Act 2022
Public Act   2022 No 39
Date of assent   8 August 2022
Commencement   see section 2
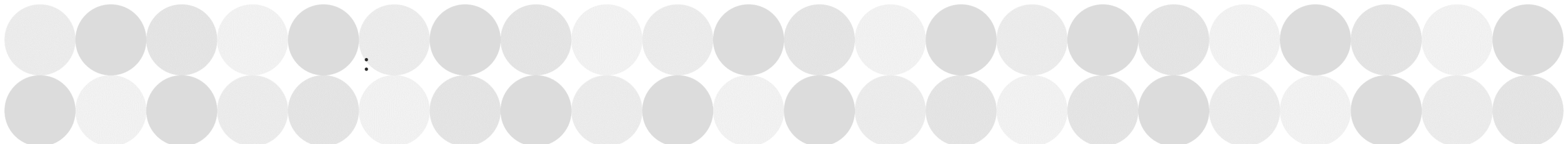
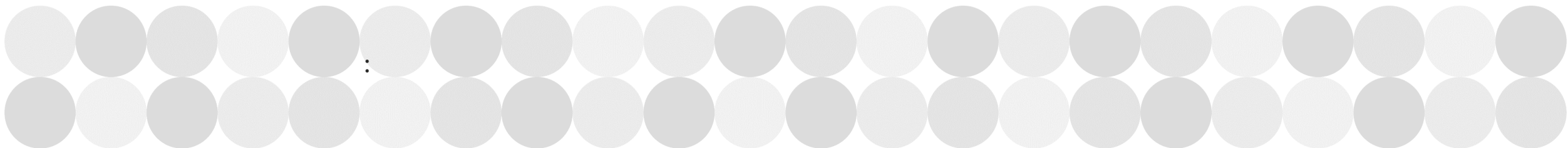# Our Tiriti o Waitangi relationship – delivering for and with Māori, Iwi & Hapū
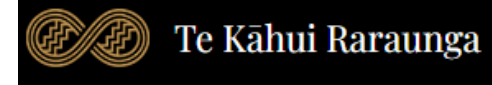
For Māori, iwi, and hapū, data is taonga (treasured/sacred)

The commitment we make to Māori:
- Progressing Māori data aspirations and enabling self-determination
- Openness to different perspectives, an informed understanding of what long-term inter-generational wellbeing looks like for Māori
- Working for and with Māori, Iwi and Hapū to improve the quality and quantity of data about Māori
- Stats NZ is investing in its people to upskill in Māori Crown relations

Ngā Tikanga Paihere – maturing our ability to apply Nga Tikanga Paihere principles to our data practices (ethical and cultural practices for data as taonga)

Te Kāhui Raraunga (Māori Data Governance Model) – Māori's expectations of the crown (Te Kāhui Raraunga)

:

Future census models

Survey **more** of the population

Survey **less** of the population

Survey first combined model with full enumeration

Admin first combined model with full enumeration

Admin first combined model with sample survey/s

Population counts and some variables from admin data + expanded social & population survey programme for other variables

2023 model

What balance of admin data and surveying should there be?

Expanded Social

Survey first + admin for gaps

Admin first + survey for some attributes

Admin first + expanded social survey programme

Readiness for admin-first

Stats NZ
Tatauranga Aotearoa

Internal readiness:

- Internal uncertainty – how do we feel about this potential change now?
- Internal barriers & resistance to change
- Change management & capability building is required
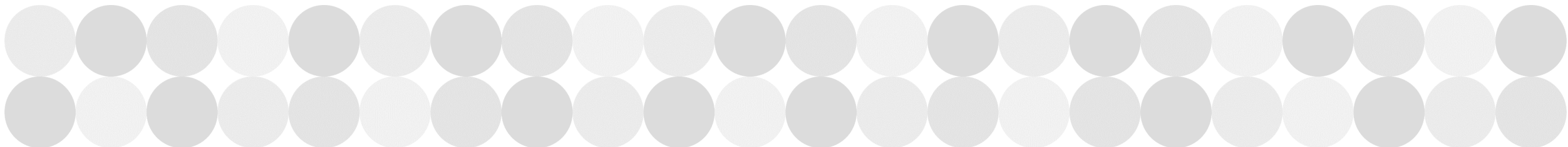- 2023 Census has taken so much of everyone's focus, next Census is hard

Methodological feasibility:

- we are well-placed from a population data perspective
- working on household & families – it is harder
- our ability to produce dwellings is less developed, people
- data sources for iwi (Māori tribe) affiliation are unresolved
- there are no major barriers

Technology Infrastructure readiness:

- to combine administrative, survey and modelled data we require linked statistical registers (SR) in a production environment outside of the IDI
- and processes to link survey & admin data to these

Our current maturity:

BUSINESS
STATISTICAL BUSINESS REGISTER IN PRODUCTION

PLACE
STATISTICAL LOCATION REGISTER IN PRODUCTION

PEOPLE
STATISTICAL PERSON REGISTER IN IDEATION

Admin-first opportunities & challenges…

Opportunities

- Frequency & detail – Outputs every year, already have an experimental 16-year time series 2006 -2021
- Quality – More precise, more detailed information, or accurate for some variables – not constrained by respondent's ability to answer questions.
- Resilience – Inherently less exposed to extreme risks or system failure.
  - For example, easier to recover from earthquakes, pandemics.
- Longitudinal – The unit record data is longitudinal – this is immensely powerful and largely untapped.
  - For example, detailed analysis of flows, cohorts, generational effects.
- Synergies with other research – Potential to be a multiplier for other research providing common baselines for researchers in the IDI.

Challenges

- Customer acceptance – how do we talk to customers in a different way and help them understand the trade-offs
- Tiriti o Waitangi obligations – how do we work with and for Māori, and meet their expectations
- Readiness for an attribute survey – how prepared do we need to be?
- Trust and confidence in use of admin data – how will we know?
- Will Stats NZ be ready – it's a huge change to population measurement and the way a census has worked!
- Authorising environment – how do we take ministers along for the journey?

Questions?