Producing Regional Profiles Based on Administrative and Statistical Data in New Zealand

Michael Slyuzberg[1] and Gail Kelly
Inland Revenue, Wellington, New Zealand
[1] Corresponding author: Michael Slyuzberg, e-mail: Michael.slyuzberg@ird.govt.nz

Abstract

New Zealand Inland Revenue recently developed 16 comprehensive regional and customer profiles which provided the agency with an in-depth view about the different characteristics of New Zealand's regions. The profiles have been used to underpin the design of resource allocation to Inland Revenue's district offices and to inform the local focus of frontline staff. A range of quantitative and qualitative data from a variety of internal and external sources was used to create the profiles, with all administrative data being geocoded during the process. This unique combination of data sources has generated new knowledge about New Zealand's regions. Currently, Inland Revenue in close collaboration with other government departments including Statistics New Zealand is working on a cross-agency initiative aimed to streamline and improve the generation of knowledge for government use. The project will synthesise data from selected agencies, create a combined regional database and apply statistical techniques to generate new knowledge from the data. This aligns with government's call for innovative use/re-use of data and information; greater openness and transparency in data use across agencies; better generation and application of knowledge to inform policy making; and better targeting of quality service delivery.

Keywords: administrative data, regional information, cross-agency collaboration, mapping and visualisation

1. Introduction
Lack of sub-national information limits the ability of government departments to develop effective regional policy. Different agencies attempt to collect such information in isolation, which is not efficient and does not produce a holistic picture of New Zealand's regions. Furthermore, many smaller agencies do not have the resources or expertise to carry out such work. The disparate, opaque holding of regional information precludes multiple use analysis and leads to poor decision making. New Zealand Inland Revenue investigated how to improve this situation.

Inland Revenue is a national tax authority playing a critical role in improving the economic and social wellbeing of New Zealanders. Inland Revenue collects 85% of the Crown's revenue as well as collecting and disbursing social support programme payments and providing the government with policy advice. Recently, Inland Revenue undertook a set of actions to enhance service delivery in order to meet increasing expectations of government and society. As part of these actions, the Research and Evaluation Unit within Inland Revenue developed 16 comprehensive regional and customer profiles which provided the agency with an in-depth view about the different characteristics of New Zealand's regions.

2. Objectives
The production of the regional profiles focussed in achieving three key objectives. First, to support and enable dynamic evidence-based decision making for multiple public sector organisations. Second, to assist the end-users, both government and non-government, to make their relevant decisions based on trusted reusable information. Finally, to increase openness and transparency of the government in line

with citizens' expectations.

In addition to these high-level objectives a number of specific, tactical goals for the initiative were established, including:

- Regular production and publication of written profiles based on a synthesised interpretation of data from selected agencies;
- Developing and supporting a combined regional database, or a number of linkable databases where full integration is not feasible;
- Developing statistical techniques to generate new regional knowledge from the data.

3. Approach

In order to achieve the above objectives, we firstly needed to develop an approach which allowed the creation of an efficient system to easily find, share and reuse high value regional information to support management decisions. It became obvious early on in the project that the scope of the product needed to be wider than a traditional database. Informing management decisions required a product that provided an interpretation of data to generate a wider, strategic regional picture based on social and economic knowledge. Therefore, we designed our product as a story telling document. Each profile consists of 60-80 pages telling a comprehensive story about a particular region.

Developing such comprehensive regional stories requires the analysis of data from multiple sources. Traditionally, the key source of sub-national data is the census of population and dwellings. However, the key disadvantage of the census data is its low periodicity – in most countries (and New Zealand is not an exception) the census is undertaken once in five years at best. The best solution in this situation is to combine the census data with other specialty surveys and publicly available administrative data. Regional profiles produced by Inland Revenue are supported by data from all of these sources.

To produce accurate sub-national stories the source data needs to be geocoded, that is, assigned to particular geographic areas. While for the census and many other survey data this may not be a problem, administrative data is not always linked with geographic units. Our approach was to geocode all source data prior to placing it to the database. In many cases it has been found that lack of geocoding prevents otherwise useful data from being included in the database or from creating a reliable time series.

Finally, before creating the product we undertook a significant amount of consultation with potential users of regional information. The purpose of these consultations was to understand what types of decisions the regional profiles would support and what data are needed to tell a comprehensive and compelling story. The discussions with the end-users proved to be an important component of the regional profiles design.

In line with the approach described above, the process of creating regional profiles included such phases as defining the structure of the stories, identifying the core data and its sources, database design, developing consistent definitions and metadata, geocoding, data cleansing and quality control, populating the database with both internal (administrative) and external data, analysis and visualisation, interpretation of the results, including local knowledge about looming regional change, writing regional stories, dissemination and analysing feedback.

4. Outcomes

The internally published set of sixteen standardised regional profiles includes the following sections:

- Overview of the region (brief history, key features that distinguish this region from others);
- Population (both overall population derived from statistical sources and the tax base population structure derived from administrative data);
- Key demographics (age, net international migration, ethnicity, education, income sources, family income levels – derived mostly from census data);
- Deprivation index (publicly available indicator of socioeconomic deprivation on a meshblock level which combines nine variables from the census);
- Housing (property sales, prices and tenure of dwelling – derived from census data and specially organised statistical surveys);
- Economy (key regional economic activities, regional GDP, industry structure, structure of businesses by size, employment by industries, building consents, business confidence – extracted from multiple publicly available external sources and, partially, from internal administrative data);
- Customer characteristics related to tax administration (customer compliance, tax debt, use of tax agents, use of electronic channels for contacting the tax authority – derived from internal administrative data).

The stories are supported by extensive visualisation (maps, graphs) and summary tables providing aggregated information about a particular region and its comparative characteristics with other regions and the country in general. Examples of visual components of the regional stories are shown below on Figures 1 and 2.
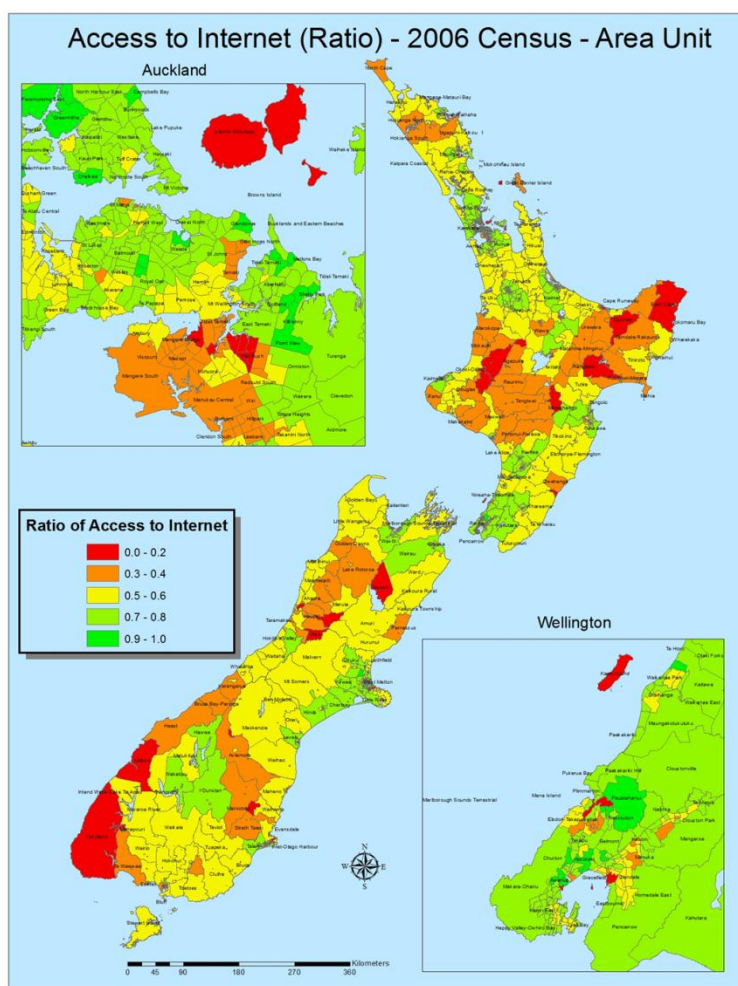

Figure 1. Access to Internet in New Zealand (ratio) on the area unit level. Source – Statistics New Zealand census data (2006).
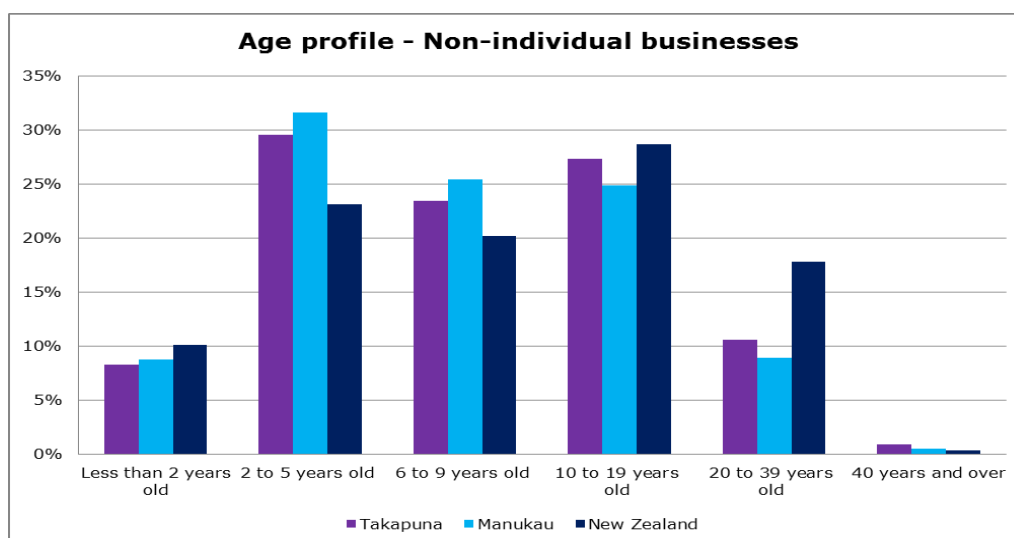
Age profile - Non-individual businesses

Figure 2. Comparison of the business age structure for non-individual (ie. excluding self-employed) businesses in the regions of Manukau, Takapuna and overall New Zealand. Source – IR administrative data.

5. Issues
The key issue in preparing regional stories was lack of timely and publicly available sub-national data from external sources. In many cases the only available source of reliable sub-national data was the census (2006) which did not reflect potentially significant changes in regional development that occurred within last six-seven years.

Further issue was lack of geocoded internal data. Although most administrative records contain, or may be linked with, relevant customer addresses, the accuracy and timeliness of these addresses is not always high and required extensive cleansing and refreshment. These efforts, unfortunately, need to be repeated periodically to maintain the quality of the regional database because of the high level of internal migration of the New Zealand population.

Use of data from multiple sources required significant work around standardisation of relevant metadata – definitions, regional boundaries and data formats. Finally, the interpretation of data, preparation of stories and dissemination of final documents required development of special skills and tools such as semi-automated reporting templates.

It is worth noting that, although the risk of breaching customer confidentiality was mitigated due to the high level of data aggregation, all stories passed a thorough confidentiality control procedure before their release.

6. Next steps
To address the key issue – lack of timely and publicly available sub-national data from external sources – Inland Revenue submitted a proposal to develop a business case for a cross-agency initiative aimed to create a whole of government regional information for decision-making. The initiative was approved and the business case development funded by the New Zealand Treasury.

The initiative is intended to provide regional information to improve policy making and to better target quality service delivery. This also aligns with government's demands for innovative use and re-use of data and information within the *Open and Transparent Government* programme.

A stakeholders' panel of representatives from seven central government agencies that have an interest in the expected outcomes or can influence the proposal has been formed. Among these stakeholders, Statistics New Zealand has a particularly important partnership role due to its key co-ordinating position in producing New Zealand's Official Statistics.

The stakeholders' panel discussed the desired outcomes from the initiative and agreed that the following characteristics would best describe these outcomes:
- Increased regional information is available to decision makers in the wider public and private sectors;
- This information is easily accessible, reducing the amount of time and effort that users need to spend finding, obtaining and re-formatting data;
- Users utilise regional information provided to inform their decision making and evaluation processes, enabling more effective analysis at national, regional and territorial levels;
- Users experience increased satisfaction in meeting their data access and information needs;
- Agencies that provide regional data and information do not need to manage separate datasets, reducing potential duplication of effort;
- Increased awareness, participation and contribution from agencies with geo-spatial interests;
- Partnerships are developed where geo-coded information obtained helps organisations identify overlapping or complementary goals and objectives; and
- Increased consistency, usability and clarity of data-sets as agencies adopt common formatting and validation of data-sets.

To support the core business case Inland Revenue is also undertaking additional activities which include a review of international best practice in collecting and publishing regional information; stocktake of regional data produced and published by government agencies in New Zealand; comprehensive public sector survey of the users of regional information; improvement of the Inland Revenue regional database to include partial automation of the extraction and quality control of internal administrative data.

It is intended to finalise all of the above activities by the end of June 2013.

7. Conclusions
The production of sixteen regional profiles supported important strategic decisions focussed on improving Inland Revenue service delivery. The key benefits from using the profiles include:
- More targeted presence within communities;
- More effective and efficient resource allocation between regional offices based on regional profiles and needs;
- Increased visibility within communities;
- Identification of specific regional issues and ability to address them;
- Developing for the first time a heat map of regional population, social demographics and tax administration issues; and
- The ability to refresh and re-run data extracts used in these regional profiles to enable a level of measurement of compliance change at intervals for each region.

The profiles also generated new knowledge. In particular, the profiles demonstrated that each region had a unique blend of people, industries and geography; and that the

"one size fits all" approach (to removing barriers to compliance, or maximising uptake of social policy entitlements) does not work.

Currently, Inland Revenue in close collaboration with other government departments including Statistics New Zealand, is working on a cross-agency initiative aimed to streamline and improve the generation of knowledge for government use. The expected outcome, in the first instance, would be written profiles interpreting synthesised data from selected agencies. This aligns with the government's call for innovative use/re-use of data and information, greater openness and transparency in data use across agencies, better generation and application of knowledge to inform policy making, and better targeting of quality service delivery.