DATA471_Group_EDA

kakauchad_300212228

2022-08-30

Background and data

In responding to COVID-19, the New Zealand Government, like all governments, took extraordinary steps to try and prevent the introduction of COVID-19 into the country (by closing New Zealand's international borders), then successfully eradicating COVID-19 with the introduction of strict controls on movement (lockdowns). After COVID re-entered the country, the government employed a range of measures to limit the spread or mitigate the impact of infections, including limits on movement, mandating the use personal protective equipment, mandating vaccinations and the use of quarantine for people arriving into the country.

This project will look at the impact of the New Zealand Government's decisions around managing the COVID-19 Pandemic (Pandemic). The analysis seeks to find relationships between key decisions made by the government and any changes in key social and economic indicators including family violence, victimisation, employment, sense of control, and public confidence in government institutions. The majority of data for this project is sourced from "StatsNZ Wellbeing data for NZ" website. (https://statisticsnz.shinyapps.io/wellbeingindicators/_w_f12c8130/?page=home).

These decisions by the government have intended consequences (i.e. managing the spread of COVID-19) and many unintended consequences were managed as well. For example, where retail closures for non-essential products and services leading to reduced income, offset by government grant and subsidies to business owners. The authors expect to identify the times where the government made key decisions and will look for changes in key indicators from official datasets:

- Personal well-being, including:
  - Institutional trust
  - sense of control

- Crime data, including:
  - Family violence
  - victimisations

- economic indicators, including:
  - employment
  - hourly rates

The objectives of this project are to link decisions to effects

The aim of this project is to determine if there is any link between the decisions made by government in managing the Pandemic and crime statistics, with particular respect to family violence and general victimisation. The main decisions of concern for this project are the decisions that changed level of constraints on personal or social interaction, including things like limiting movement (by way of lockdowns), limiting social interaction (restrictions on group size), and mandating additional safeguards (.e.g use of personal protective equipment (PPE), vaccination, self-isolation etc.). Intuitively, those decisions should have obvious effects on two economic indicators (i.e. the employment rate as businesses struggle to retain staff during periods of low activity; and subsequently, income levels as the proportion of revenue generating activity reduces due to government imposed limitations), and one well-being indicator (i.e. sense of control, as the government makes decisions that limit what citizens can do).
As the Pandemic continued, the initial tide of public support for stamping out COVID-19 waned as the impact of the different restrictions started to accumulate in more far-reaching ways, which should result in an effect on a second well-being indicator (i.e. trust in parliament) as citizens recalibrate their views on whether the state is acting within its own powers and in the best interests of its citizens.
I hypothesise a less direct linkage between government decisions and two crime statistics, family violence and victimisation. The underlying assumption for the first linkage is that family violence is connected to and driven by an offender's sense of control and exacerbated by increasing pressures (from all quarters, but

in this context, potentially reduced income and reduced sense of control). The expectation is that as sense of control decreases *and* income decreases, the prevalence of victimisation will increase generally, and the prevalence of family violence will increase specifically.

To that end, the key objectives of this project are: - identify key decision points by government in responding to the Pandemic, and - define the essential nature of the decision into one of four categories: - increased restriction - decreased restriction - increased economic support - decreased economic support - describe well-being indicators, including: - visualise and describe how sense of control changes - visualise and describe how trust in parliament changes - describe economic indicators, including: - visualise and describe how employment rate changes - visualise and describe how income changes - describe crime indicators, including: - visualise and describe how family violence changes - visualise and describe how victimisation changes - identify relationships between the decisions and any of the indicators, including: - correlation between indicators - indicators of lead or lag between indicators

Data characteristics

The datasets are drawn from official data, including the Household Labour Force Survey (supplementary), Police statistics, and the Economic Survey. As official datasets, each set comprises multiple categorical variables identifying aspects such as series, sub-series, indicator-names, and parameters. Each set also includes numerical variables for the main value of interest, and most include a time-stamp, most often as a categoric marker for the period which the data represents.

Plots of key indicators

The key indicators in this project are 'sense_control', 'employment_rate', 'income', 'trust_in_parliament', 'family_violence', and 'victimisation'. The following is a simple representation of summary data available on some of the key indicators. The data is aggregated to an annual level and one of the biggest challenges with these datasets is the inconsistent aggregations and depth of data. For instance, economic data is collected weekly whereas the well-being indicators were collected annually, and more recently on a quarterly basis. Similarly, economic data has been collected for decades, whereas well-being data has been collected for a shorter time period. In order to provide a consistent basis, the data has been aggregated to an annual level and data for only 2018-2021 is used (for consistency with the family violence and victimisation data). 1 shows the column plots of the aggregated data.

Note that the total family violence variable peaked in 2019 and began to trend downwards for both 2020 and 2021. This same pattern occurs for total victimisations, with a peak in 2019 and slight reductions over the following two years. Total hourly rate increased year on year from 2018 to 2021 and the total employment rate was fairly steady throughout the period, although there was a small decrease from the start of 2020, before the employment rate began to lift again from the middle of the year. the employment rate continued to rise through 2020 and remained on the rise the start of 2022, when it stagnated.

Ethical considerations

This project looks at the effects of government decisions on society, by considering a range of statistics, many gathered for a specific purpose (e.g. the salary: hourly wage rate is collected as an economic indicator, trust in public institutions is collected as measure of confidence in pillars of society etc.) but will be used as proxies (individually or in combination) in this project for estimating some other aspect of society. While the information used is publicly available, we need to be aware that project output may have effects beyond those intended from the original collection. One of the key assumptions in this project is that a driver of victimisation and family violence is income, more accurately, that as income decreases, we expect to see an increase in at leasto one of those indicators. Perhaps most concerning from an ethical position is that
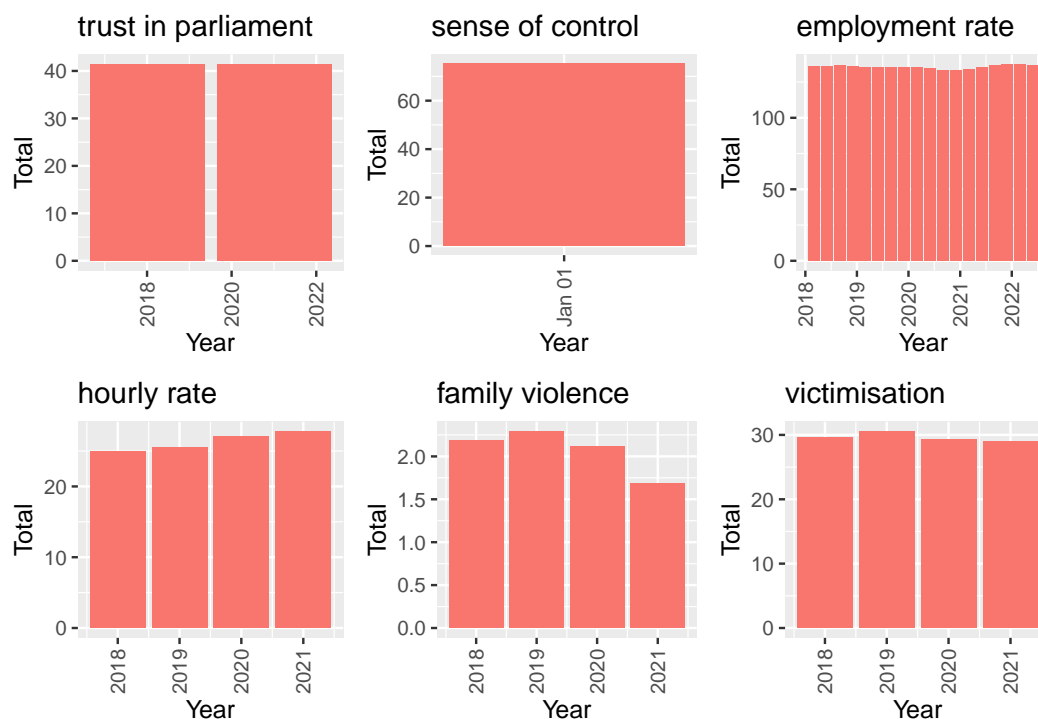
Figure 1: Column plots showing annual totals of the six key indicators under study: trust in parliament, hourly rate, sense of control, family violence, employment rate and victimisation

family violence is expected to increase as income rate and sense of control decrease, which has the risk of stigmatising or stereotyping lower income groups as more likely to commit family violence.

Security considerations

Drawing from official datasets, there are very few security considerations with this project, since the underlying data is held on secure, government owned servers, with only portions of the data available for access by the public. More granular detail is available on request from StatsNZ, but even then, access to the underlying data itself is unlikely and much of the data released is in a summarised form (i.e. rather than allow users to access data directly, StatsNZ produces summary data and makes that available for users). This document is held on a publicly accessible github repository and will remain accessible for some weeks after conclusion of the project before the repository is made private in 2023. The underlying data used in this project is also available at that same github repo, but can be accessed directly from the StatsNZ Infoshare website. No additional security requirements beyond github's built in security mechanisms are used. Device level controls are used for coding, data and work-up of this document.

Privacy considerations

The initial dataset from StatsNZ Wellbeing indicators is large and contains several indicators, that are quite revealing in terms of the themes that are represented and collected from survey participants and official sources. Although there is possibility that the data in the raw dataset could be aggregated to identify individuals or specific groups, this project takes very high-level approach (i.e. annual, quarterly, or monthly

level aggregations) to using the data. Identification of individuals or specific groups is unlikely at the level data is used in this project.

Government decisions - timelines and decision types

We extracted text data from the McGuinness Institute website (https://www.mcguinnessinstitute.org/projects/pandemic-nz/covid-19-timeline/) and used that as the basis for establishing a decision timeline. Decisions are categorised into two types: 'restriction' and 'economic support'. Restrictions are decisions about some constraint on private individuals or organisations and can take the form of restrictions of movement (e.g. border closure, lockdown, limit on the number of people in a space, mandatory vaccinations or isolation etc.); or economic support (e.g. provision of subsidies, relief payment, government funding etc.). For restrictions, an increase occurs when the decision results in a higher level of restriction (i.e. reduction in freedoms) and a decrease occurs when the decision results in greater freedoms. Similarly, increases in economic are those decisions that result in additional funding or government spend or a reduction in compliance costs, and a decrease in economic support is where funding is reduced, or announcements result in private individuals or the private sector bearing the costs of the decision. The McGuinnes Institute data includes announcements and summary information, so although there are over three hundred datapoints, not everyone will qualify as a decision point.

Initial analysis of the McGuinness data shows concentrations of particular terms which are shown in Fig **??**. The plot is quite busy and covers a range of announcements that aren't clearly related to restrictions or economic support so we can filter to some of the more obvious choices, including filtering by announcement category (e.g. Alert level, Bubble, Traffic light, Financial support, and Governance).

Scatterplot of announcement types, by date
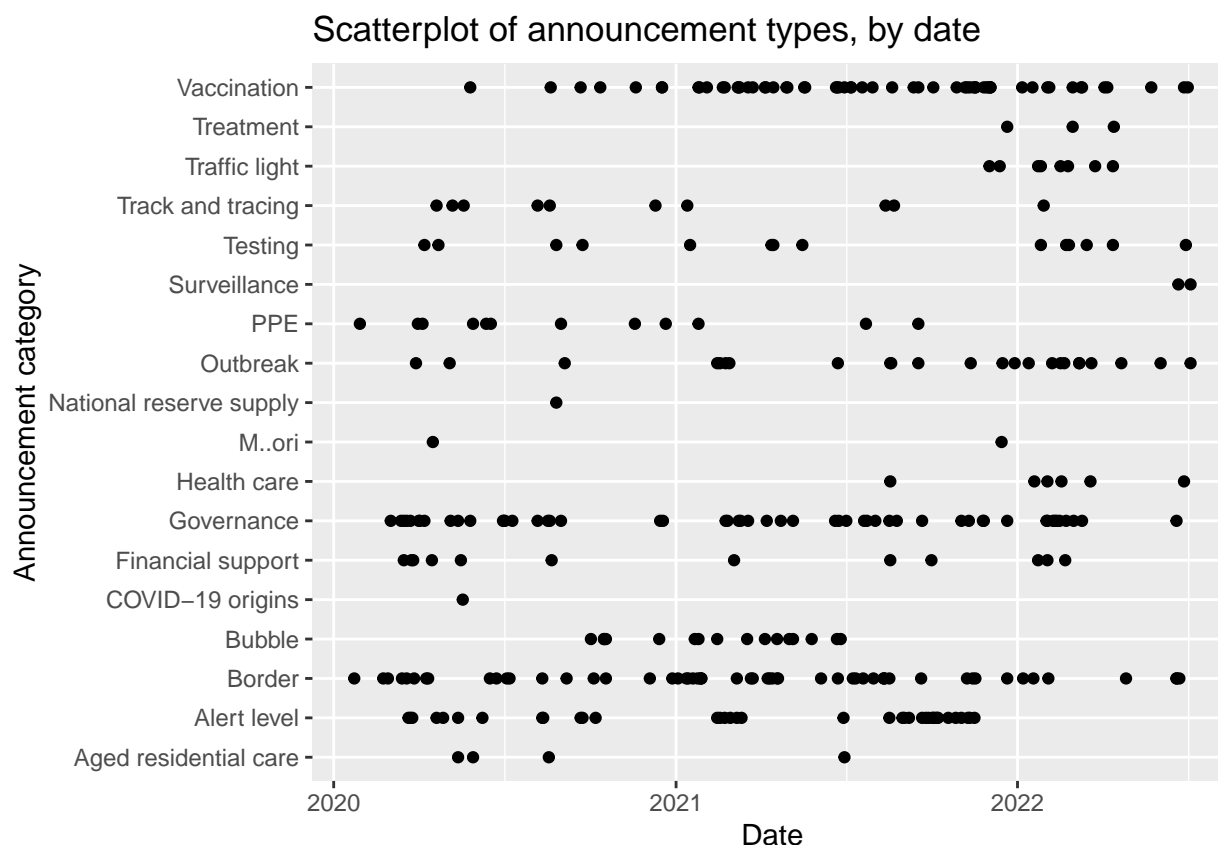
Alert levels, bubbles and traffic lights

The scatter plot shows a decrease in the frequency of announcements about Alert in the fourth quarter of 2020 and a shift to more announcements in the Bubble category. In the first quarter of 2021 there is a high frequency of announcements about alert levels, before a reduction in the frequency of those announcements to almost none for the 2nd quarter. In the third and fourth quarter of 2021 there are no more announcements about bubbles but, again a high frequency of announcements about alert levels. Announcements about Traffic light system begin only in late 2021, at the same time that announcements about alert levels cease completely.

Financial support

Announcements about Financial support are fairly rare through the course of the Pandemic, beginning in early 2020 at about the same time as announcements about alert levels and about Governance. In fact, every announcement about financial support occurs at about almost the same time as an announcement about Governance and often an announcement about alert levels, most prominently through 2020 (and for the six financial support announcements through 2021 and 2022). Notably, when the three announcements about financial support were made in 2021, there is a high frequency of announcements about alert levels, and for the three financial support announcements in 2022, a correlating announcement on traffic lights.

Vaccinations, governance, and border

Vaccinations are not mentioned much until the third quarter of 2020 and then mentioned very frequently throughout 2021, with perhaps only Governance and Border announcements at the same level of frequency.

Announcements about vaccinations are made frequently, and the most consistently throughout 2021, and slightly less frequently in the first quarter of 2022. Vaccines announcements are made more frequently from quarter 1 of 2021, including when very high frequency of announcements about alert levels are made in the first and third quarters of 2021.

Announcements about Governance occurred st frequently during the second and third quarters of 2020, then again throughout 2021 and tended initially accompanied announcements about the border through 2020 and also accompanying announcements about vaccinations through 2021. There is a very high frequency of announcements about governance in the first quarter of 2022.

What can we take away from this?

Even without reading the detail of each announcement, the scatterplot provides a sense of the types of decisions being made by government: Announcements about alert levels are accompanied by announcements about financial support, likely an effort to soften the economic impact and improve compliance, as well as announcements about governance, likely to both maintain public confidence and to provide clarity about how and why decisions were being made, especially in the early part of the pandemic.

Similarly, announcements shifting between alert levels, bubbles, and later, towards traffic lights, indicate the changing methods for managing risk. Again, where announcements about alert levels increase, there is also an increase in the frequency of governance and financial support announcements.

As vaccines became more available worldwide, and subsequently in New Zealand, announcements about vaccination remained frequent, right up to early 2022, when fewer announcements were made, especially at the same time that governance announcements became much more frequent. This period coincides with the occupation of Parliament grounds by anti-vaccine and anti-mandate protesters.

economic and well-being indicators

Although there were significant announcements made to impact on both economic and well-being indicators the summarised data tends to point to improvement in many of those indicators - family violence and victimisations decreased over the 2019- 2021 period, hourly earning rates went up and so did employment rates over the same period. The acknowledgement has to be made that the economic and indicator data was summarised to a national level, and annualised in most cases. It would be interesting to see how those indicators change if slicing the data by sector (e.g. retail, hospitality, health-care etc.) and by region.

project improvements

There are three areas this study could improve: 1. stricter approach to defining government decisions, 2. maintaining granular data for a longer through the analytic process to establish any key patterns 3. improve analysis of text data, both of the data sourced during this analysis as well as of any other supplementary data

defining government decisions

In this analysis, government decisions were not really defined, they were loosely inferred from announcement categories. At the most basic level a government announcement should have a definition, most likely against an official source, noting that many decisions were policy level decisions rather than statutory. Statutory decisions are fairly easy to track but policy decisions often involve a reallocation of existing resources to achieve some different but desirable end. The actual decision may be more public relations than public good.

analysing granular data

This analysis summarised data early, partly to reduce the data overheads, and partly to be able to make slightly more meaningful comparisons between datasets, driven by the different depths and frequency of collection. The key slice, for economic and well-being indicators, would be an examination of relationships between regional groupings and the indicators - does the East Coast have a different profile than South Auckland, or Wellington City etc. One of the interesting aspects of working in an urban environment was seeing how the adoption of work-from-home practices saw an increase in the number of people (and employers) recognising that without the burden of physical presence, the job could still be done just as well... with fewer personal costs (e.g. transport). How many people left the cities to go to rural (or even semi-rural) climes?

improving text analysis

One of the things I wanted to do was do some text analysis to see if I could discern any language patterns in the different announcements as a way of checking the tone of the message, for the atmosphere of the day.

conclusion

In conclusion, this was a pretty interesting project that highlighted a couple of surprise outcomes from how the New Zealand government handled the pandemic, including increase employment rates and reduced family violence rates during a time of unprecedented change and innovation occurred.