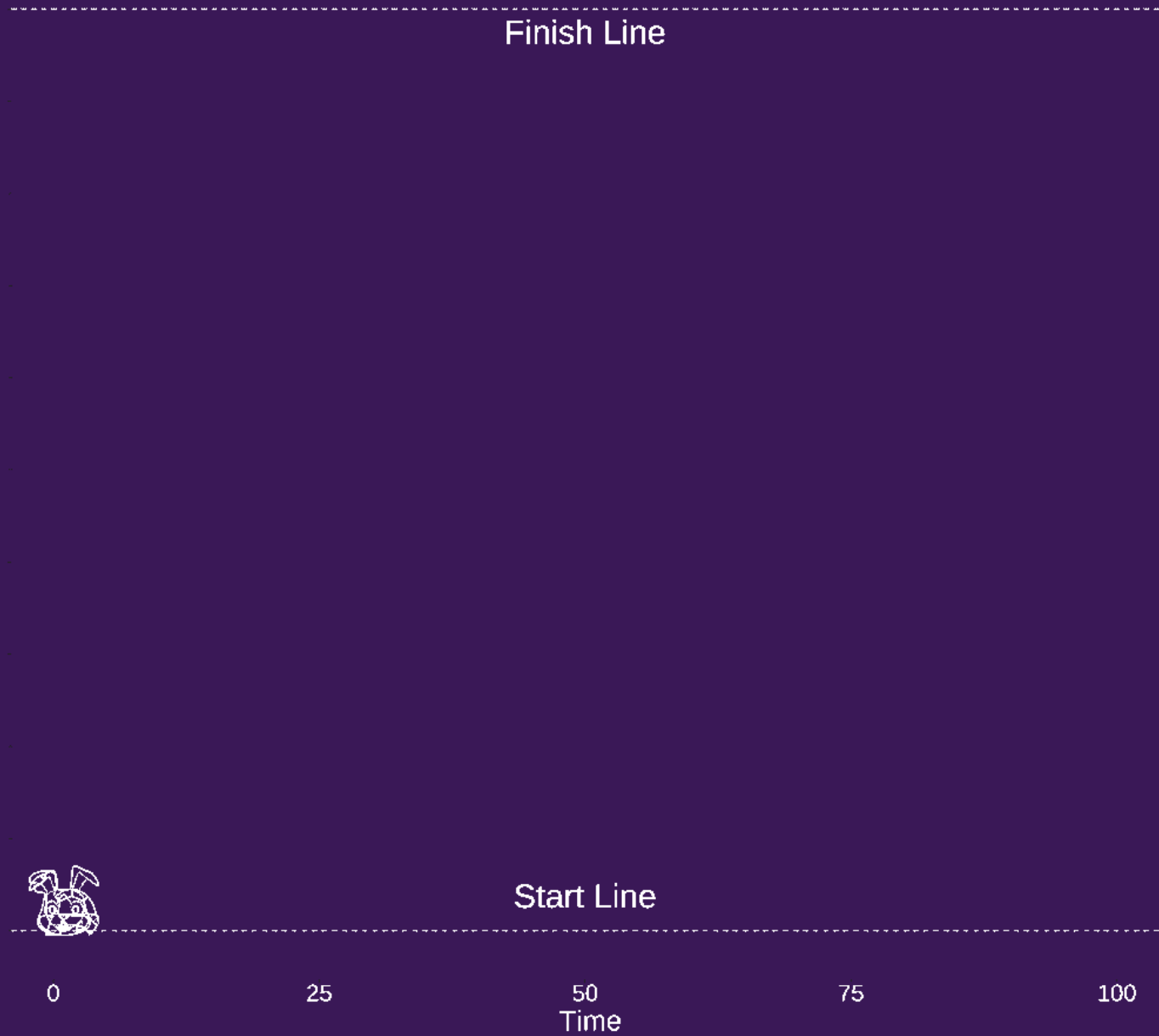


Tortoise and Hare Race



Data Storytelling with R and ggplot

Dr Kevin Tsang
Associate Professor

2 July 2026





**Storytelling
element 1:**

Audience

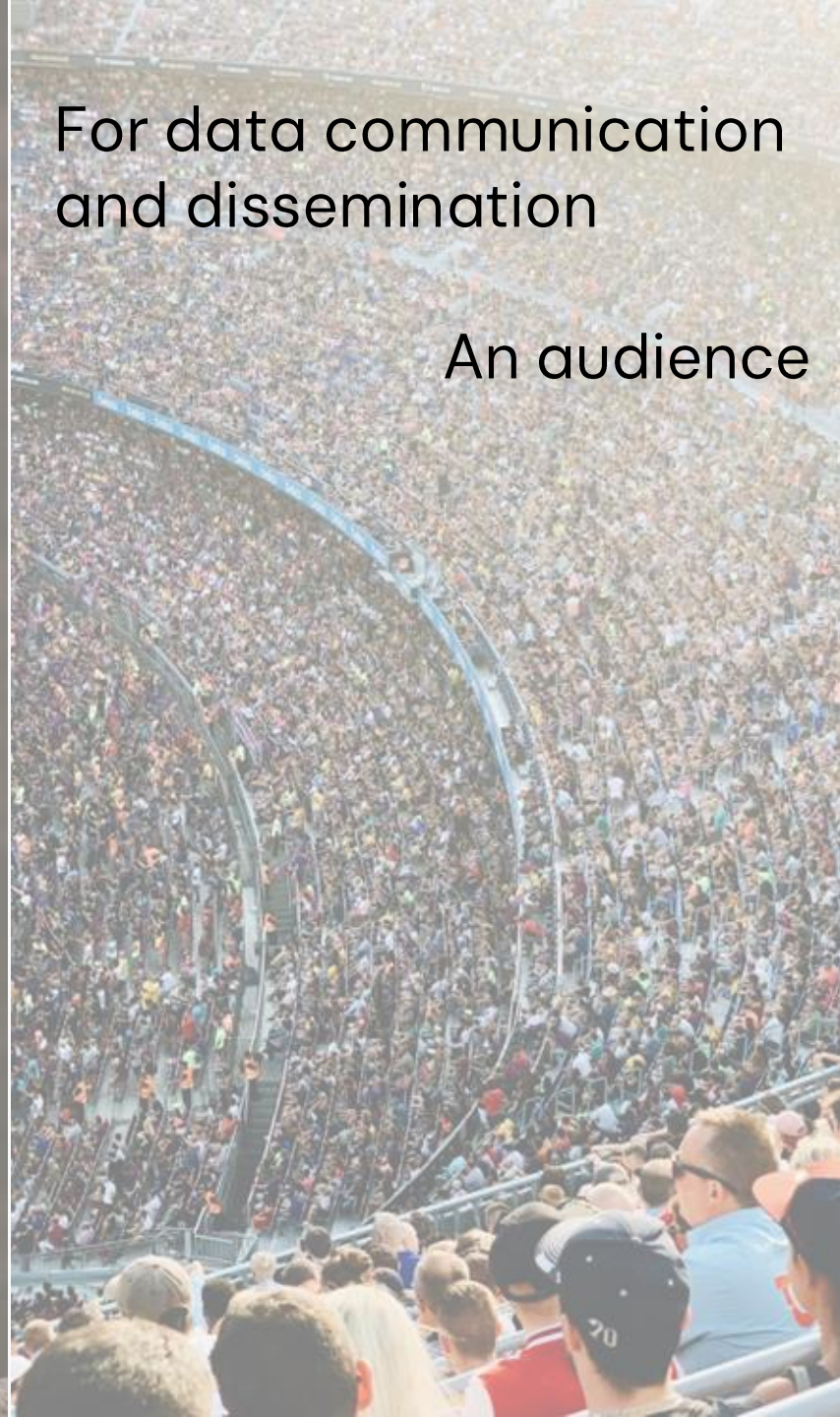
For data analysis
and exploration

Yourself
Your team



For data communication
and dissemination

An audience



For art

Yourself
An audience



www.data-to-art.com

Adapted from: Julian Hoffmann

5 ways to capture an audience

5

Number



Picture



Movement

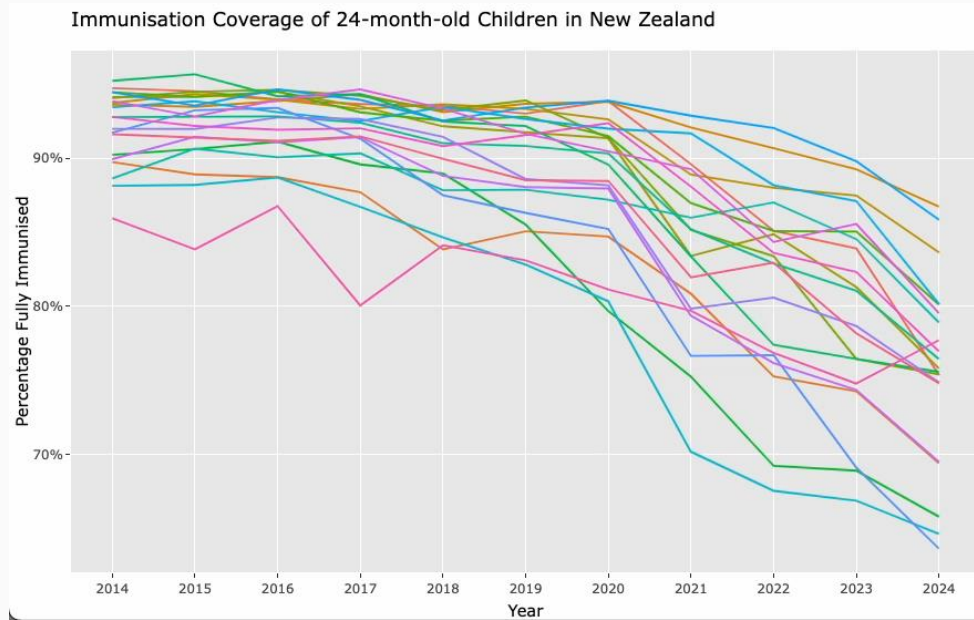
“Any intelligent fool can make things bigger and more complex. It takes a touch of genius - and a lot of courage - to move in the opposite direction”

Quote

Tragedy of Darth Plagueis the Wise

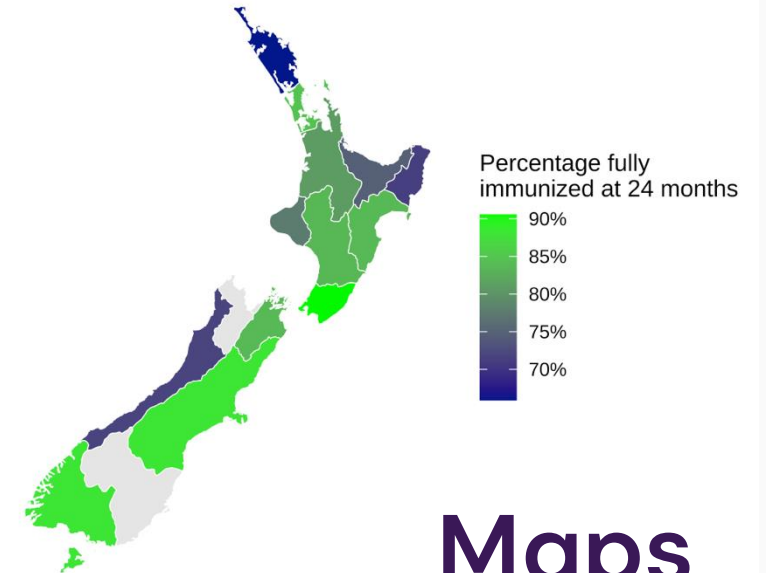


Story



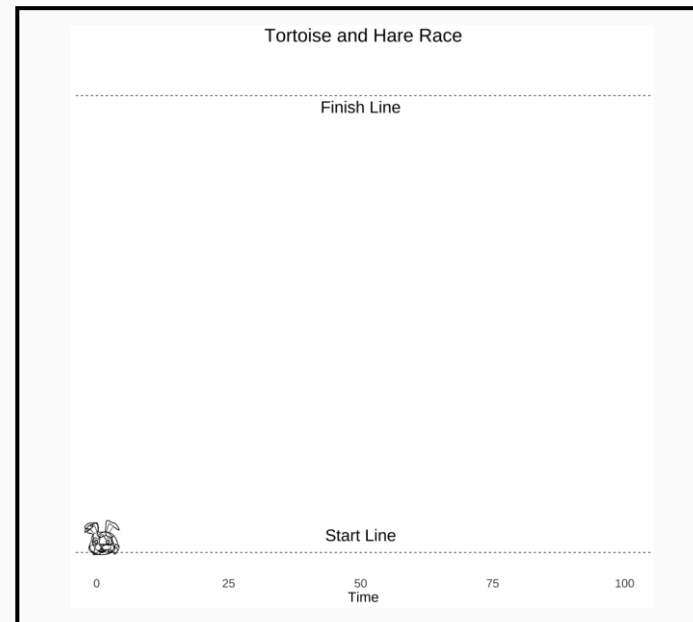
Childhood Immunization Rates by Region (2026)

DHBs aggregated to regions



Data: Ministry of Health NZ

Interactive



Maps

Animated





Storytelling
element 2:

Setting

Data description

Immunisation coverage of children at 24 months old in New Zealand

1 January 2015 – 31 March 2026

<https://www.healthnz.govt.nz/about-us/health-data/data-sets-and-collections/immunisation-data-and-statistics/immunisation-coverage>

The 5 health targets



Faster cancer treatment

90% of patients to receive cancer management within 31 days of the decision to treat.



Improved immunisation for tamariki

95% of tamariki fully immunised at 24 months of age.



Shorter stays in emergency departments

95% of patients to be admitted, discharged or transferred from an emergency department within 6 hours.



Shorter wait times for first specialist assessment

95% of patients wait less than 4 months for a first specialist assessment.



Shorter wait times for elective treatment

95% of patients wait less than 4 months for elective treatment.

Data exploration (for you)



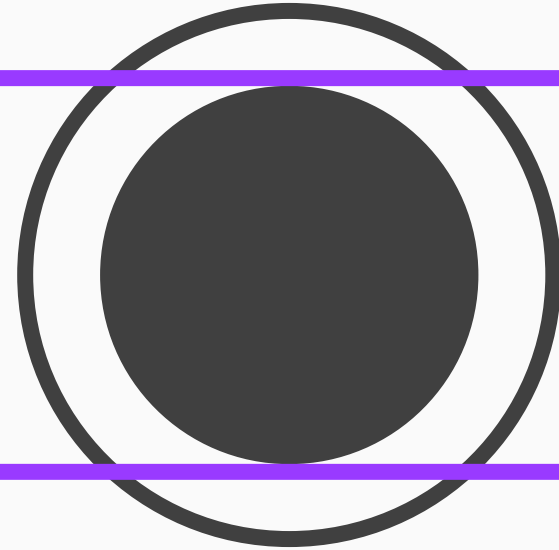
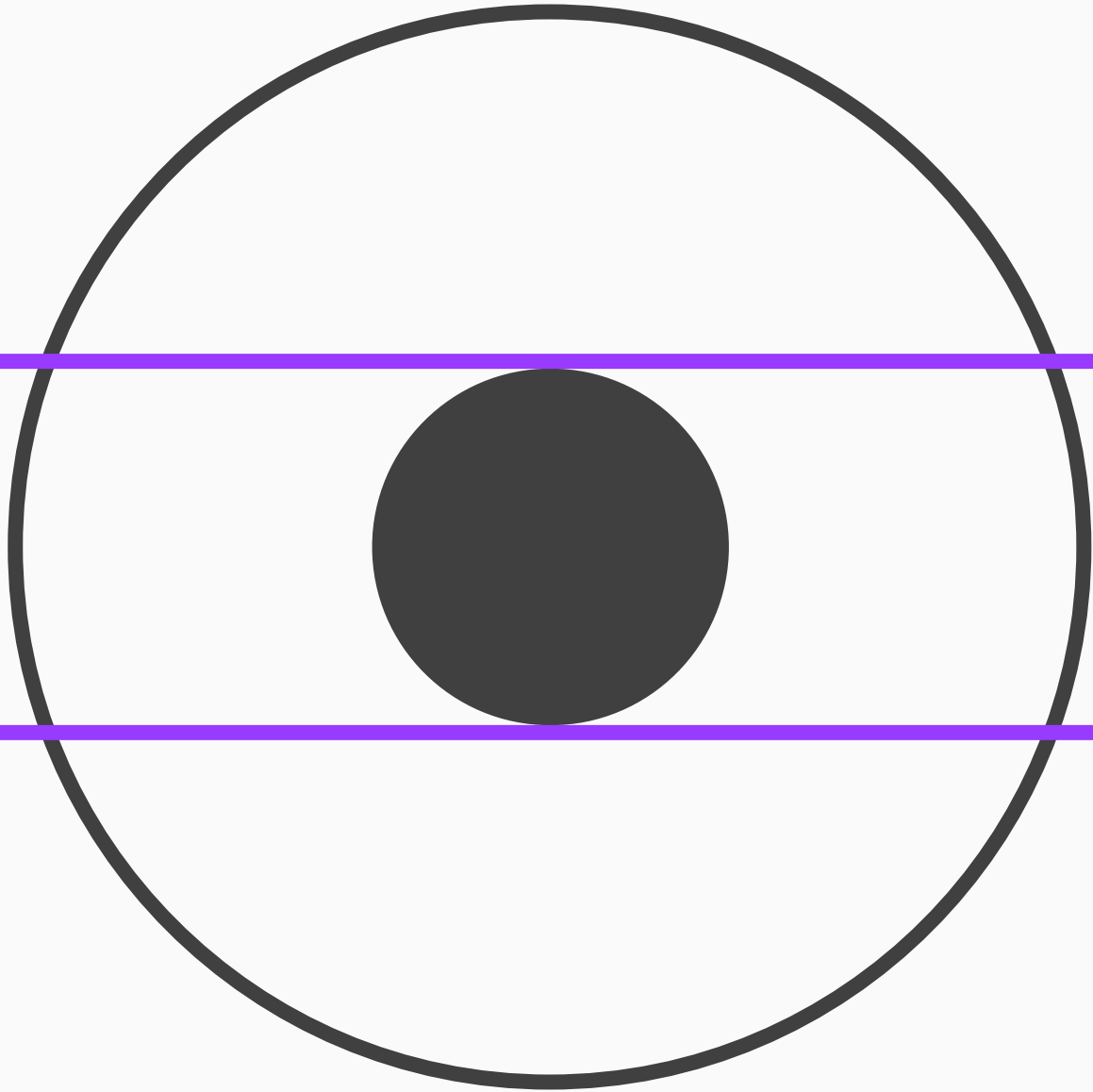
Formulate questions



Identify patterns

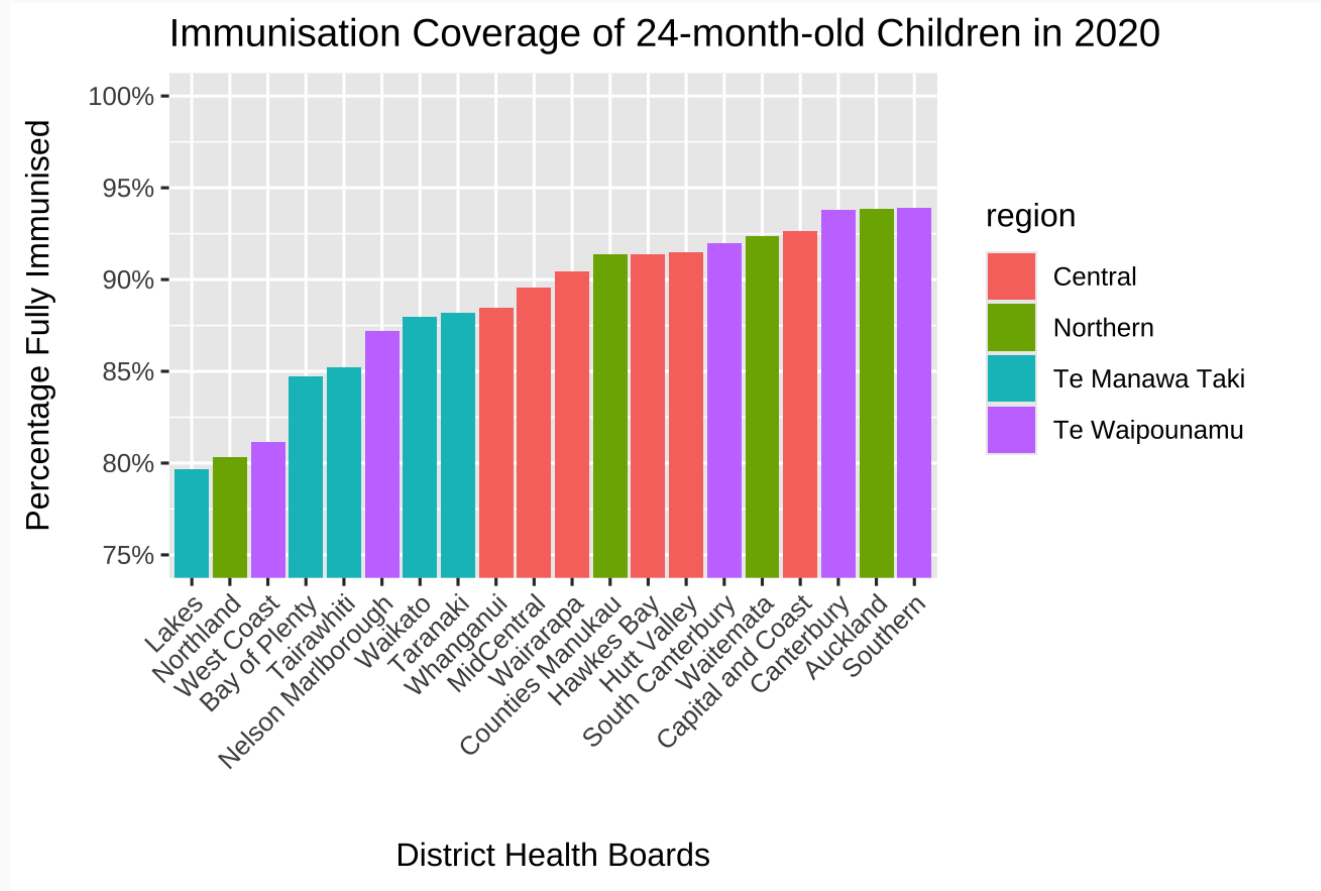


Understand the landscape



Explore 3 aspects of the data with 1 plot

Dimension 2:
y-axis



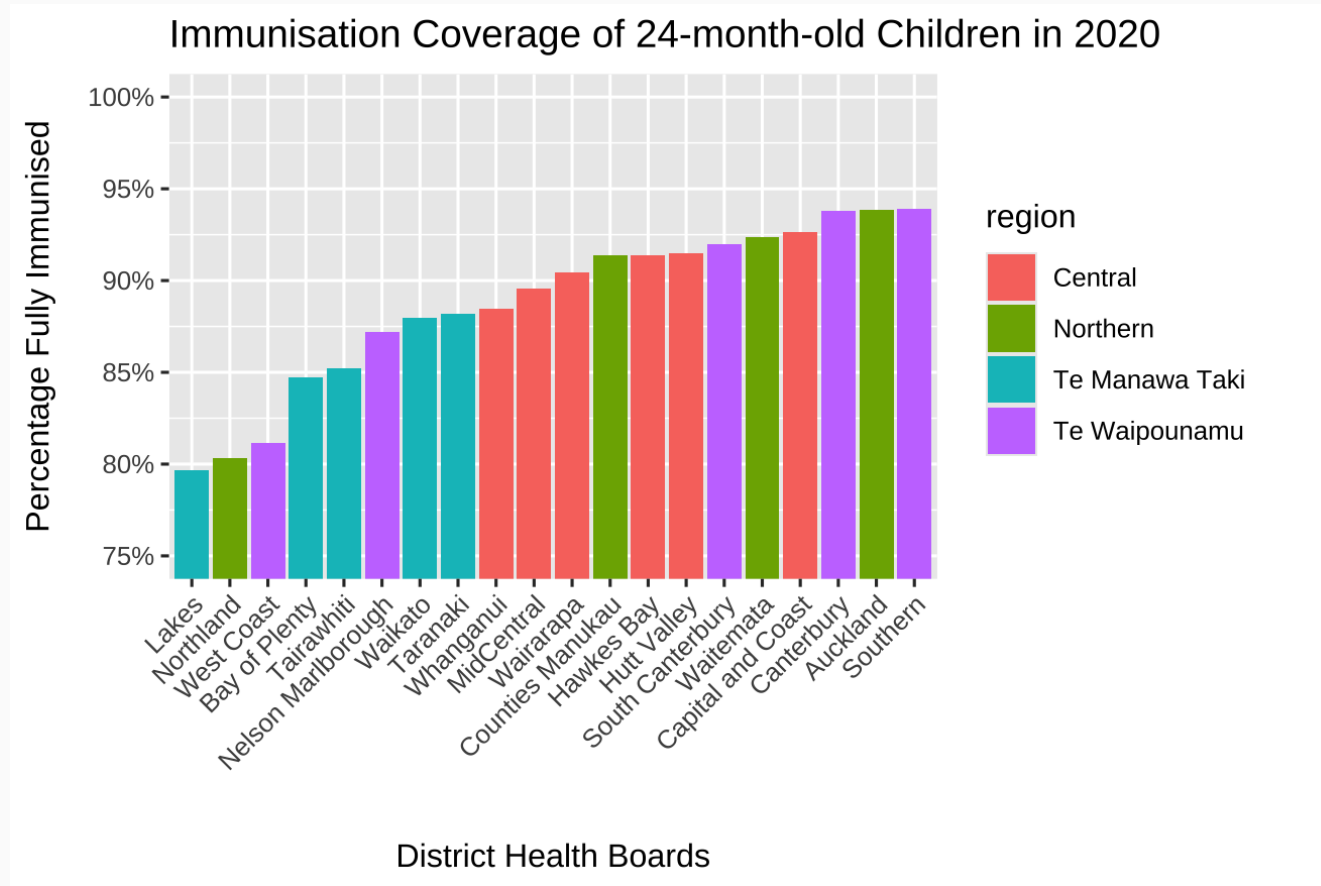
Dimension 3:
colour

Dimension 1:
x-axis

Explore the 3 aspects of the data with 1 plot



What is the highest and lowest coverage across NZ?



Are there any links between coverage and region?

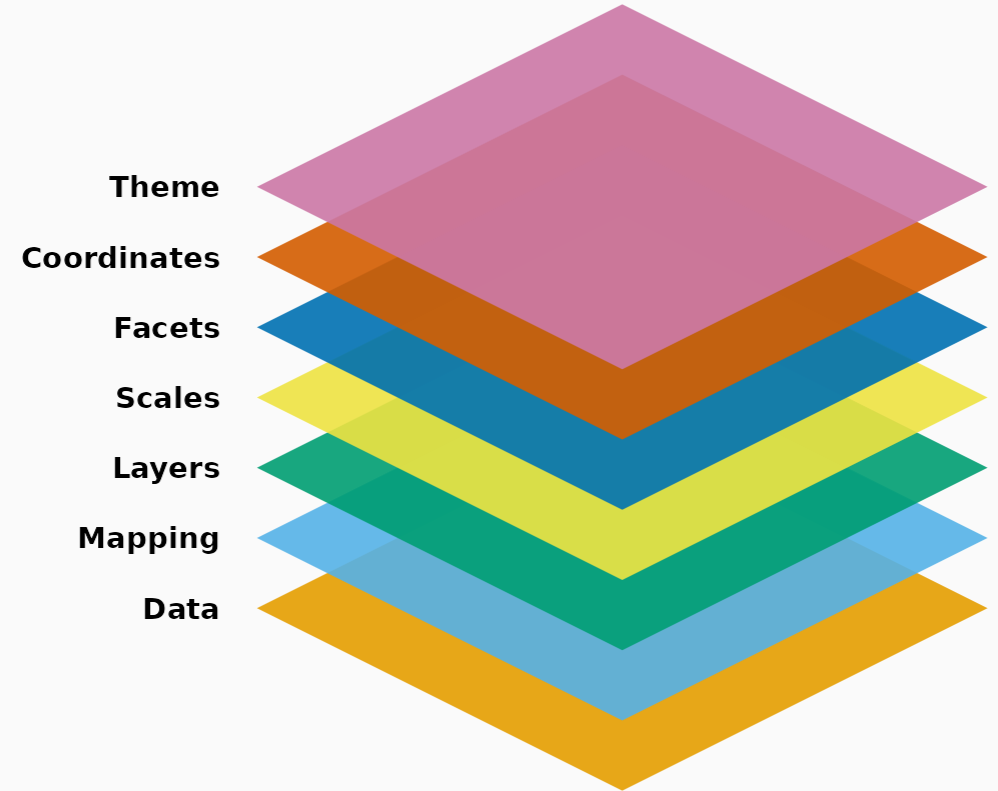


What is the coverage in my district?



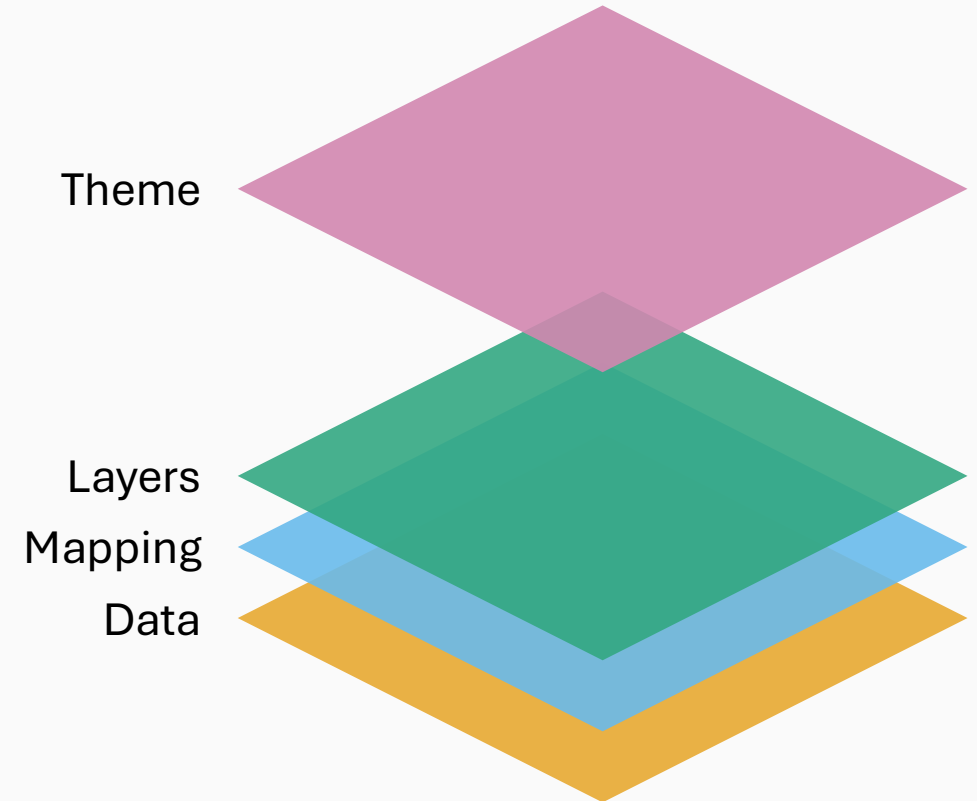
ggplot2

```
ggplot(<data>,  
  aes(<mapping>)  
) +  
<layers_function>() +  
<scales_function>() +  
<facets>() +  
<coord_function>() +  
<theme_function>()
```



ggplot2

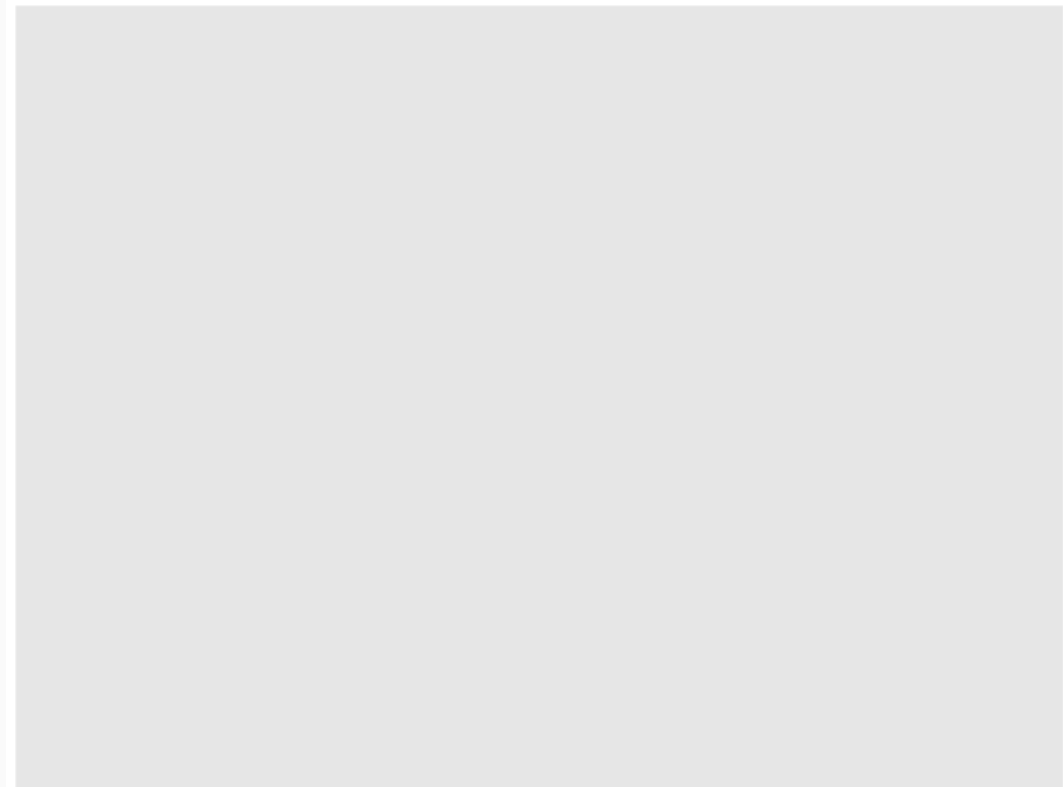
```
ggplot(<data>,  
      aes(<mapping>)  
    ) +  
  <geom_function>() +  
  <theme_function>()
```



ggplot2

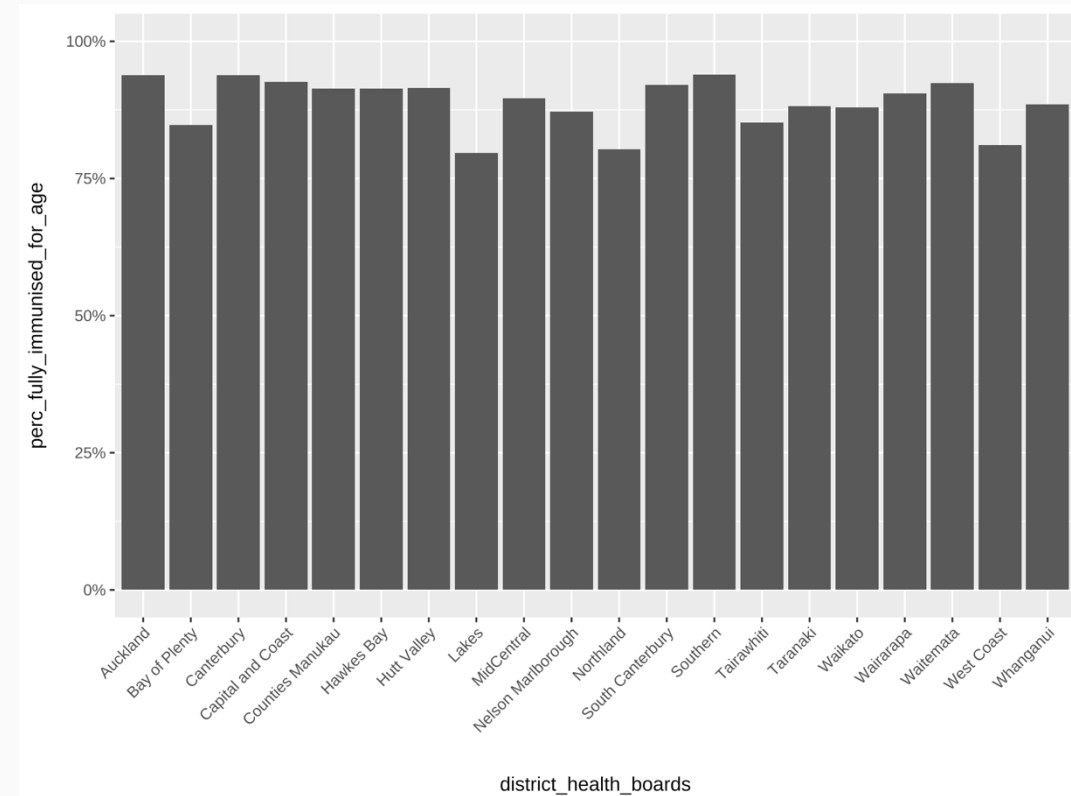
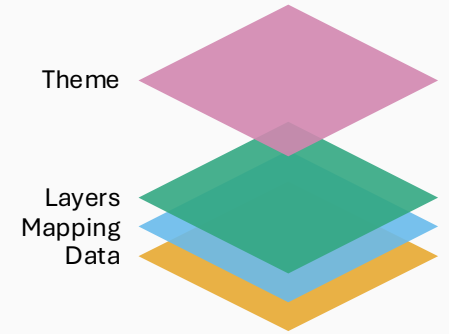
```
ggplot(child_imm_2020_dhb)
```

Data



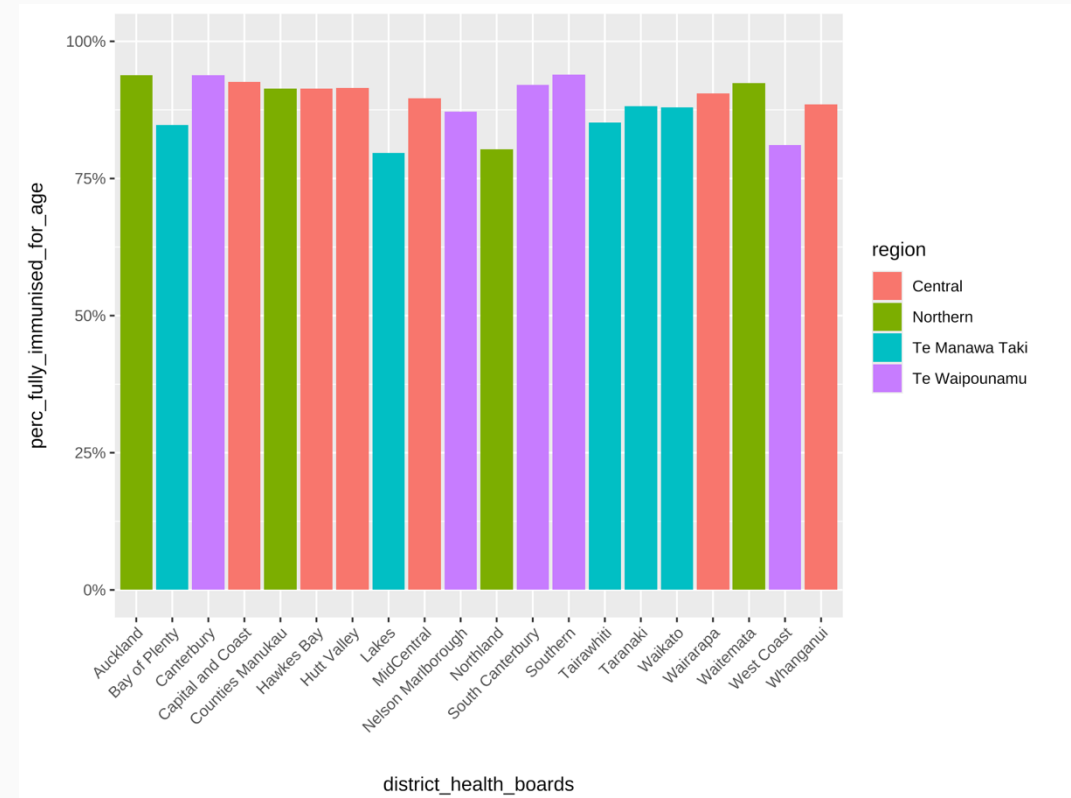
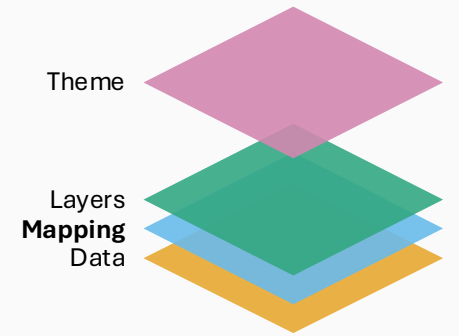
ggplot2

```
ggplot(child_imm_2020_dhb,  
  aes(x = district_health_boards,  
      y = perc_fully_immunised_for_age))  
geom_col() +  
theme(axis.text.x = element_text(  
  angle = 45,  
  hjust = 1))
```



ggplot2

```
ggplot(child_imm_2020_dhb,  
  aes(x = district_health_boards,  
      y = perc_fully_immunised_for_age,  
      fill = region)) +  
  geom_col() +  
  theme(axis.text.x = element_text(  
    angle = 45,  
    hjust = 1))
```

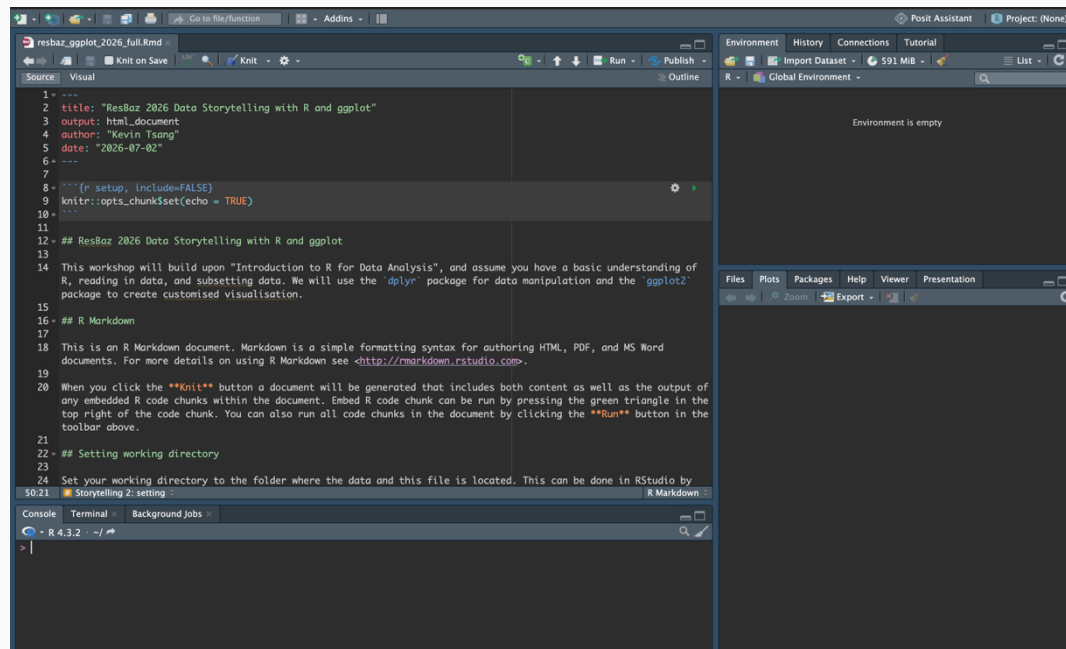


To RStudio...

You will need to download 2 files:

resbaz_ggplot_2026.Rmd

childhood_24month_immunisation
_2015_2026.csv



year	district_health_boards	region	ethnicity	num_eligible	fully_immunised_for_age
2015	Auckland	Northern	Total	6003	5675
2015	Auckland	Northern	Maori	737	693
2015	Auckland	Northern	Pacific	1087	1061
2015	Auckland	Northern	Asian	1793	1707
2015	Auckland	Northern	European or Other	2386	2214
2015	Bay of Plenty	Te Manawa Taki	Total	3002	2669
2015	Bay of Plenty	Te Manawa Taki	Maori	1255	1128
2015	Bay of Plenty	Te Manawa Taki	Pacific	78	69
2015	Bay of Plenty	Te Manawa Taki	Asian	210	203
2015	Bay of Plenty	Te Manawa Taki	European or Other	1459	1269
2015	Canterbury	Te Waipounamu	Total	6111	5712
2015	Canterbury	Te Waipounamu	Maori	944	872
2015	Canterbury	Te Waipounamu	Pacific	314	301
2015	Canterbury	Te Waipounamu	Asian	784	747
2015	Canterbury	Te Waipounamu	European or Other	4069	3792
2015	Capital and Coast	Central	Total	3617	3412
2015	Capital and Coast	Central	Maori	695	649
2015	Capital and Coast	Central	Pacific	391	380



**Storytelling
element 3:**

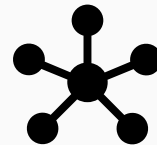
Characters

Who or what is the story about

(Who would the audience be interested by?)

For example:

One data point –
one district health board
experiencing a change in policy



Group of data points –
data points from one region

A derived variable –
year-on-year difference
in coverage



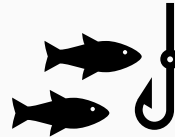
Who else is in the story

For example:



Characters on the same journey

Characters who changes main character's trajectory



Characters who reveal new dimensions of the main character



Storytelling element 4: Message

A title should capture the message

Don't be afraid to use text elements like subtitles

Annotations guides the reader

Look here



Captions can be used to add additional context or information

What actions does your data visualisation inform?



Audience



Setting



Characters

What stories will
you tell with
data?

✉ kevin.tsang@ucl.ac.uk

 [kcht](#)



Message