STAT3001 Data Visualisation Project

Pavel Misiun (20579072), Peter Misiun (20886060), Yutathkarn Coles (20579988), Ryan Sheik (20972312)

1. Problem statement

The recent global pandemic caused by coronavirus disease 2019 (COVID-19) has been described as the most consequential global health crisis since the 1918 influenza pandemic, with at least 7,000,000 reported deaths as of writing (1). The initial catastrophic impacts of COVID-19 prompted the fastest development and global deployment of emergency vaccines in history, with vaccines developed by Pfizer, moderna, and astrazeneca approved for clinical use as early as December of 2020 (2). In Australia, following the first reported cases in 2020, a swift response to COVID-19 was initiated through the implementation of various public health measures to thwart the propagation of disease, resulting in some of the best clinical outcomes globally (3). Despite breakthroughs in clinical understanding of the disease, outbreaks continue to occur, and scepticism about the response to COVID-19 prevail, with 10-30% of Australian adults who are reluctant or refuse to be vaccinated (4). The proposed report aims to explore key features and impacts of the widespread adoption of vaccines in Australia through data visualisation.

2. Research questions

Key questions will be addressed, such as:

1. The rates of vaccination, and how rapidly they were deployed relative to the emergence of COVID-19 in Australia.
2. The primary impacts of widespread vaccine deployment on disease cases, recoveries, hospitalisations, and deaths.
3. The secondary impacts of vaccines on socio-economic metrics, such as individuals' day to day life and GDP.

To answer the first research question, we can use graphs in order to see when the pandemic began within Australia and how rapidly the vaccination initiative was put into place. For the second question, we can plot how many people were getting vaccinated, disease cases, recoveries, hospitalisations, and deaths. From here we can interpret the data in order to see how the vaccination effort impacted the other factors.To answer the third question, we could plot changes in mobility and changes in GDP, interpreting this data in relation to the vaccination effort plot.

3. The data sources you are planning to use for the project.

We will use various datasets that may represent key features of interest. First, we will use the Coronavirus pandemic country profile for Australia from Our World In Data (5), which contains various covid relevant data from the start of the pandemic to present. A secondary dataset used will be the COVID-19 Open Data Repository (6), specifically, the aggregated dataset filtered for Australian data. This contains data only between 1/1/2020 and 17/9/2022, but provides

interesting insights into mobility, cancelled public events, and government spending. Finally, we will use the Australian National Accounts dataset from the ABS (7) to investigate changes in metrics such as GDP, focusing from the start of the pandemic to present.

4. Brief literature review - this should include a brief overview on the topic of the project, and if relevant, on use of data visualisation in the relevant literature to address questions in that domain.

The COVID-19 pandemic has greatly impacted Australia since it began, affecting health and daily life. The first case within Australia was reported on January 25, 2020, serving as the start of the impacts observed within Australia. Throughout the year of 2020, the most notable impact seen were the 905 deaths that occurred due to the virus (8). Another notable impact of COVID-19 was its economic impact within Australia. This is shown as for 2020, the GDP fell by approximately 2.4% and the percentage of unemployed individuals increased by 47% (8). Australia's strategy in order to combat the negative effects of COVID-19 focused on eliminating the virus. This included strict travel controls and local lockdowns during outbreaks (9). In order to combat these outbreaks, vaccination efforts started in early 2021 and as a result of this, over 70% of adults were fully vaccinated by October 2021 (10). The implementation of data visualisation tools assisted in tracking the vaccination progress with real-time data, leading to an easier understanding of impacts of implementing vaccinations within Australia (8).

With this data being recorded, it has shown that the implementation of vaccines within Australia correlates to less cases of COVID-19 and less severe symptoms for those infected by the virus (8). In addition to this, it is seen that higher vaccination rates also correlates to positive effects within the Australian economy. This is seen through a rise in GDP and lower unemployment rates, as multiple affected sectors (retail, tourism etc) were able to recover from the restrictions imposed (6). These positive impacts were influenced by data visualisations, in which they assisted in illustrating these trends for policymakers (8).

In summary, COVID-19 has had a significant impact on Australia’s health system and economy. The implementation of vaccinations and strict health measures have shown to be vital within Australia in terms of controlling the virus. However, even with these measures being implemented, ongoing monitoring will still be required in order to get a better picture at the long term effects on society due to COVID-19.

5. Interpretation of Graphs created in R in relation to the COVID-19 impact on Australia

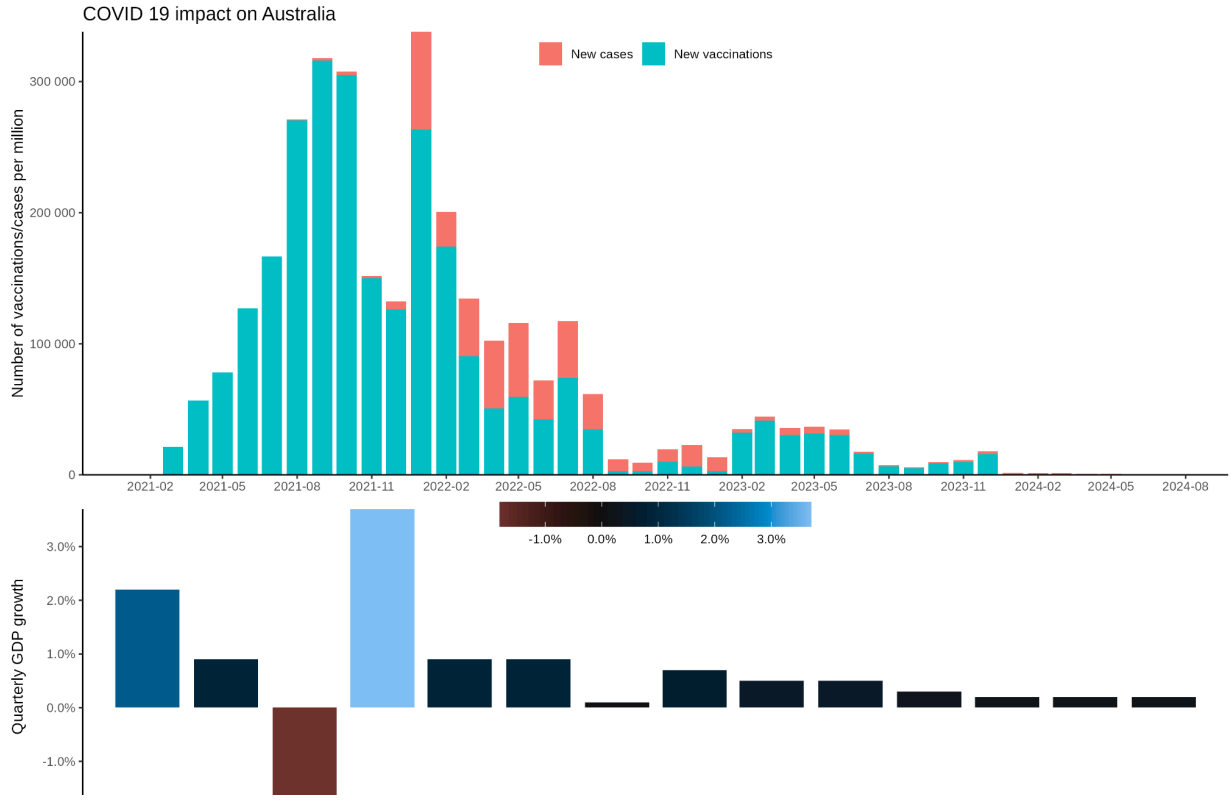

Figure 1, stacked bar chart (new COVID-19 cases in proportion to the number of Australians receiving vaccinations) and bar chart (GDP growth)

Upon conducting exploratory analysis using the aggregated COVID-19 dataset, we have gained further insight into the number of new COVID-19 cases on a quarterly basis from 2021 until the present day. A stacked bar chart highlights the relationship between new COVID-19 cases in proportion to the number of Australians receiving vaccinations, measured in hundreds of thousands.

The first quarter of 2021 displayed a relatively small number of cases as the global COVID-19 pandemic was in its initial stages. Worldwide attention and fears of this disease sparked concerns among the global population, resulting in the development and widespread use of vaccinations as a preventative measure. Our stacked bar-chart illustrates that almost all new COVID-19 cases in this period were proportionally matched by the number of Australians taking a COVID-19 vaccination. The subsequent three quarters saw strong vaccination periods, with the number of new cases being matched by vaccinations. A notable spike in new cases occurred in the third quarter, alongside an upturn in vaccinations, highlighting one of the most significant impacts on the Australian healthcare system during this period.

The end of 2021 marked the most pertinent time for the Australian healthcare system, with the total number of cases peaking at over 300,000 new COVID-19 cases reported in the last quarter. Interestingly, the rate of new vaccinations decreased during this same period, likely due to a multitude of facets. Many Australians had already received the required vaccinations for work and other essential purposes, along with growing sentiment about the vaccine's effectiveness among a minority of the population.

The first quarter of 2022 saw a significant reduction in new cases to just under 200,000, a decrease in excess of 100,000. However, the number of new vaccinations also declined as restrictions eased and Australia gradually shifted to a post-lockdown phase. The subsequent three quarters of 2022 saw fluctuations in COVID-19 cases, ultimately dropping well below 100,000 new cases in the last quarter. This decline represents the effectiveness of measures such as lockdowns and mask-wearing in public to control the spread of the virus. Unfortunately, 2022 also saw a further decline in the ratio of new vaccinations relative to new COVID-19 cases on a quarterly basis.

Interestingly, 2023 saw relatively stagnant numbers of COVID-19 cases across all four quarters, with mid-2023 experiencing the lowest rates nationwide, sinking to approximately a couple thousand cases. This trend signifies the post-peak pandemic phase, with minimal increase at the start of the year partly responsible due to effective measures such as lockdowns and workplace access restrictions. The ratio of new vaccinations in relation to new COVID-19 cases saw an upturn during this period. Data for 2024 is only available for the first quarter, showing a minimal number of reported cases, making it difficult to measure the effectiveness of new vaccinations in controlling the pandemic.

In conjunction with analysing COVID-19 cases and vaccination rates, we supplemented this analysis with an examination of Australia's quarterly GDP growth to assess the pandemic's effect on the economy. The first quarter of 2021 saw a 2% growth in Australian GDP, likely because the COVID-19 pandemic had not fully impacted the economy at that point. Subsequent periods revealed a faltering economy, ultimately seeing a shrinkage of more than 1% in GDP. Remarkably, the period following this contraction saw a dramatic upturn, with a staggering 3% growth within the quarter. This growth can be attributed to Australia's improving control over the pandemic's effects, combined with government subsidy payments to citizens and businesses aimed at promoting greater economic activity.

2023 and 2024 experienced quarterly stagnant periods, with a gradual trend towards slower growth rates entering 2024, falling below 1% per quarter. It's important to note that mid-2022 saw minimal to no growth, which can be partly attributed to the peak number of COVID-19 cases recorded during this period

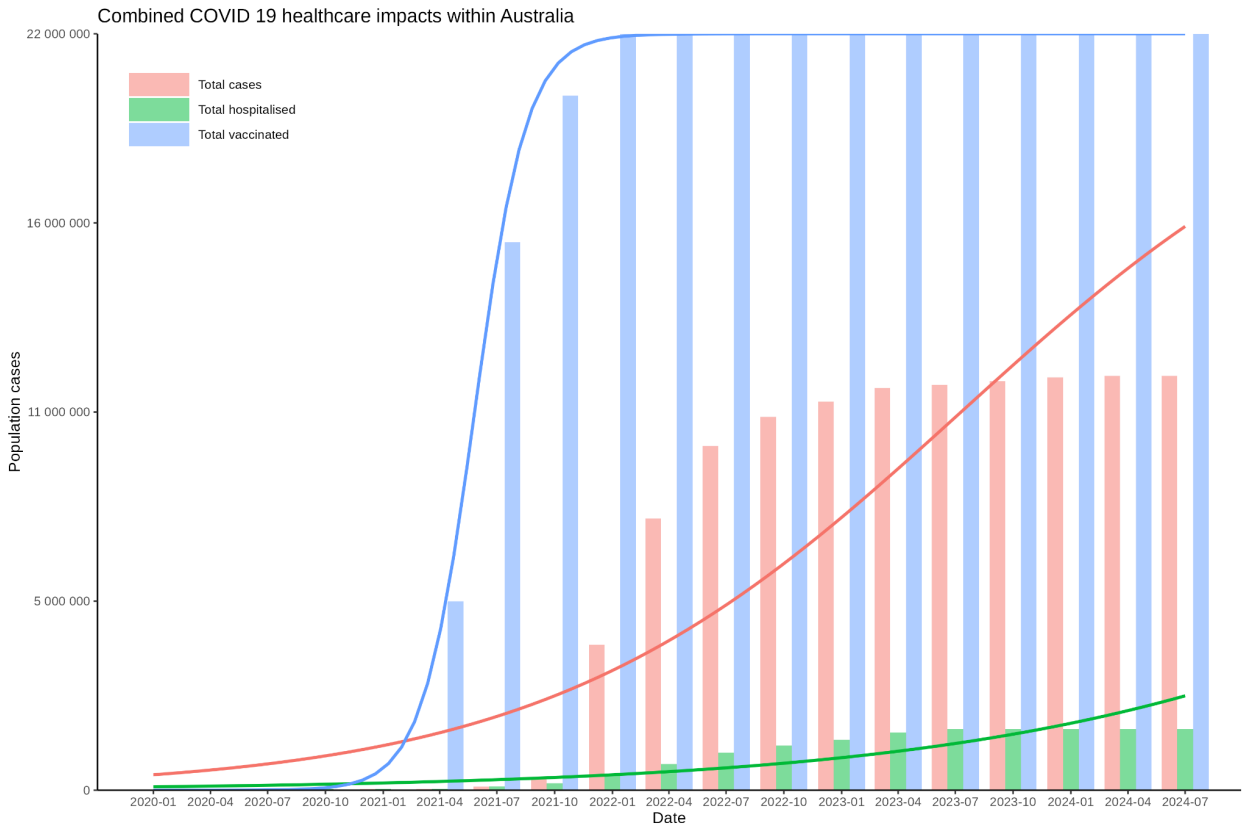

Figure 2, combined cumulative chart (total number of COVID-19 cases, vaccination rates, and hospitalisation rates)

The combined cumulative chart highlights three key measures proportional to each other: total number of COVID-19 cases, vaccination rates, and hospitalisation rates. The trend lines for these measures provide a simplified yet insightful analysis of the COVID-19 data. The blue trend line, representing total vaccination rates, shows an increase specifically between the second quarter of 2021 and the first quarter of 2022, followed by stagnation until the third quarter of 2024. This stagnation represents how a majority of Australians have received COVID-19 vaccinations, leaving only a small minority unvaccinated.

Similarly, the red and green trend lines are logistically fitted. The red line, representing the number of COVID-19 cases, shows major growth nationwide between the first quarter of 2022 and the first quarter of 2023, followed by stagnation, signifying the decline in the rate of growth and effective management strategies. The green trend line, representing hospitalised patients, saw a smaller yet significant proportional rise from the second quarter of 2022 (the peak of new COVID-19 cases) until the first quarter of 2023, despite being a smaller statistic compared to vaccination rates and total case numbers.

6. Dashboard created in PowerBI and interpretation of results

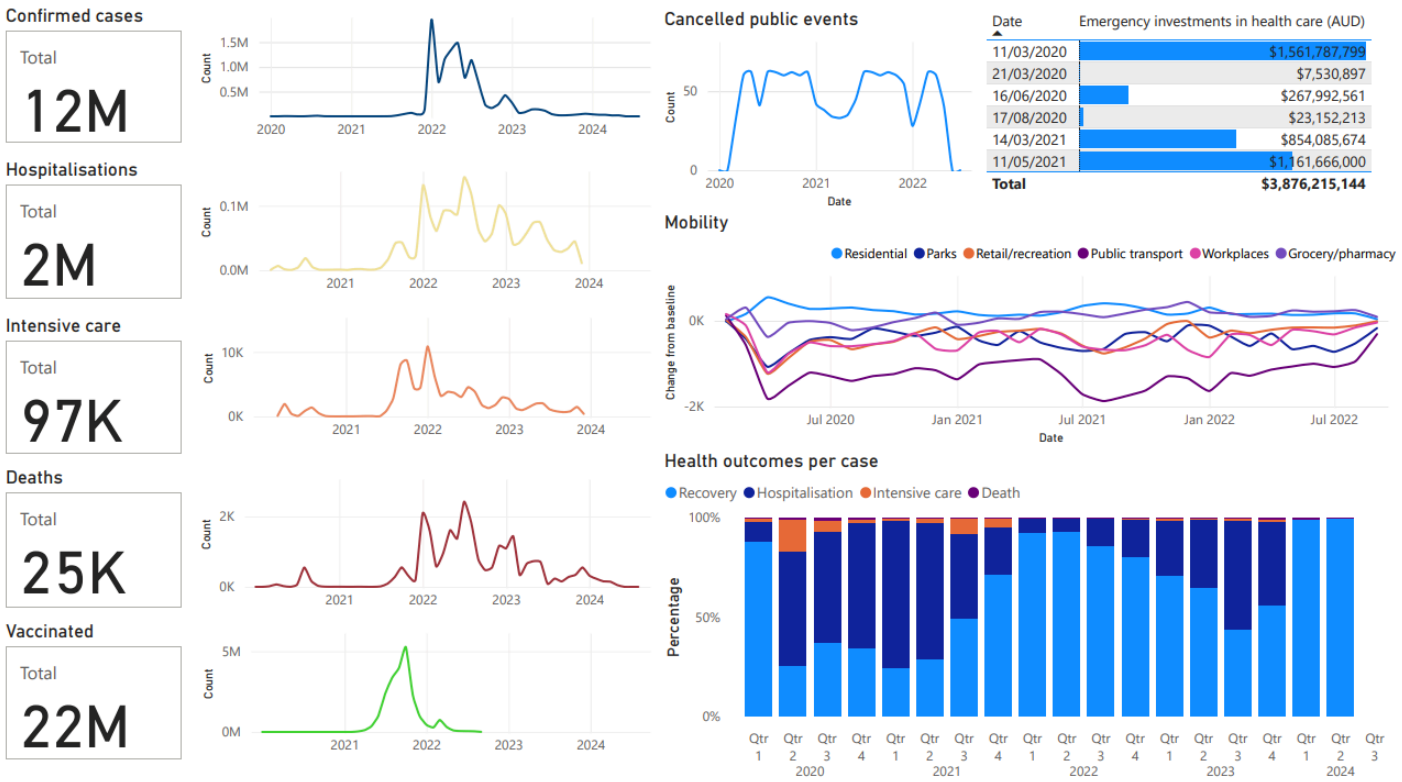

Figure 3, Dashboard displaying COVID-19 data for Australia

The dashboard above offers insights into the rapid vaccine deployment and its impact within Australia, providing answers to our primary research questions.

On the left side of the dashboard, line graphs track COVID-19 trends, including total confirmed cases, hospitalizations, intensive care admissions, deaths, and vaccinations. Throughout 2020, minor spikes in these figures represent the initial emergence of COVID-19 in Australia. Also, the stacked bar graph shows that in the second quarter of 2020, an increase in hospitalizations and intensive care admissions were present, reflecting the virus's initial impacts.

The visuals in the top-right highlight cancelled public events (line graph), emergency healthcare investments (table), and mobility in high-traffic areas (multi-series line graph). A major spike in event cancellations (from 0 to over 50) and an initial healthcare investment (about $1.5 billion AUD) coincide with a mobility drop in parks, retail, transit, and workplaces, as lockdowns and social distancing measures were implemented to limit the propagation of disease between individuals. In contrast, mobility in residential and essential retail spaces remained relatively stable, with increases in residential mobility from baseline. This is expected as these spaces, particularly groceries and pharmacies, are essential to basic needs, whereas parks, transit, and workplaces are less essential and were subject to lock-down policies. These trends represent

the Australian government's initial response to combat the impacts of the virus, leading to the social impact of limiting activity in public spaces.

Following this initial period, cases dropped, leading to reduced intensive care admissions, less event cancellations, and a slight upturn in the mobility throughout public spaces, a trend that continued until mid-2021. Around this time, the rollout of vaccines began, represented by the green line graph on the bottom left. This shows a steady monthly rise and a peak at about 5 million individuals being vaccinated monthly in the third quarter of 2021; eventually resulting in 22 million vaccinated individuals as of present. This rapid vaccination initiative was supported by two additional healthcare investments (approximately $2 billion AUD in total) in the second quarter of 2021, displaying the rapid response put in place as millions were vaccinated monthly.

Coinciding with peak vaccinations, a significant increase in cases impacted Australia between late 2021 and early 2023, with intensive care rates going from ~2,000 to 10,000. However, by the latter half of this year, the stacked bar chart displays a decline in proportions of intensive care admissions and deaths. Though the vaccine did not prevent the outbreaks seen throughout 2022, it is clear that it was effective for reducing severe symptoms caused by the virus, alleviating strain on the healthcare system. Early 2022 also saw the inception of the Omnicron variant of Covid, which, despite being less virulent, was far more contagious than previous variants, explaining the surge in cases (11). We also see that the proportions of hospitalisations and intensive care admissions slowly rise throughout 2022 and early 2023. Given that the vast majority of full vaccinations occured in mid 2021, this trend may represent a gradual weaning of the efficacy of the vaccine. Despite this, re-administration of vaccine doses were not necessary beyond 2023, with significant declines in reported cases.

7. Conclusion

The combination of visualisations effectively addresses all aspects of our research questions. The stacked bar chart illustrates the relationship between vaccination rates and new COVID-19 cases on a quarterly basis, providing clear insight into our first research question on the rate and timing of vaccine deployment relative to COVID-19 spread. The cumulative column chart further highlights the combined effect of total recorded COVID-19 cases in relation to vaccinations and hospitalizations, addressing our second question regarding the impact of widespread vaccine deployment on hospitalisation rates. Finally, the bar chart representing GDP growth reveals the secondary economic impacts of the pandemic. This combined with the dashboard's mobility and cancelled event data responds to our third research question regarding the socio-economic consequences linked to vaccination and recovery phases.

Collectively, these visualisations offer a comprehensive view of the pandemic's multifaceted effects on both public health and the economy, illustrating a dynamic picture of how Australia responded to, managed, and gradually stabilised the outcomes of disease. The improvements observed in the first and second quarters of 2024, as depicted on the stacked bar chart (dashboard), show a marked reduction in hospitalisations with a majority of cases progressing to recovery. This outcome highlights the potential benefits of ongoing vaccination efforts and a

shift toward preventative healthcare measures. Together, these trends reinforce the importance of proactive public health strategies and provide a data-driven basis for future health policy, supporting a shift towards resilience and preparedness for potential future outbreaks.

References

1. COVID-19 deaths | WHO COVID-19 dashboard [Internet]. [cited 2024 Sep 29]. Available from: https://data.who.int/dashboards/covid19/deaths
2. Ledford H, Cyranoski D, Van Noorden R. The UK has approved a COVID vaccine — here's what scientists now want to know. Nature. 2020 Dec 3;588(7837):205–6.
3. Basseal JM, Bennett CM, Collignon P, Currie BJ, Durrheim DN, Leask J, et al. Key lessons from the COVID-19 public health response in Australia. The Lancet Regional Health – Western Pacific [Internet]. 2023 Jan 1 [cited 2024 Sep 29];30. Available from: https://www.thelancet.com/journals/lanwpc/article/PIIS2666-6065(22)00231-0/fulltext
4. Tran MH. COVID-19 vaccination hesitancy in Australia: a public health issue. Front Public Health [Internet]. 2024 Jan 18 [cited 2024 Sep 29];12. Available from: https://www.frontiersin.org/journals/public-health/articles/10.3389/fpubh.2024.1282986/full
5. Australia: Coronavirus Pandemic Country Profile - Our World in Data [Internet]. [cited 2024 Nov 3]. Available from: https://ourworldindata.org/coronavirus/country/australia
6. COVID-19 Open Data — Google Health [Internet]. [cited 2024 Sep 29]. Available from: https://health.google.com/COVID-19/open-data/raw-data
7. Australian National Accounts: National Income, Expenditure and Product, June 2024 | Australian Bureau of Statistics [Internet]. [cited 2024 Nov 3]. Available from: https://www.abs.gov.au/statistics/economy/national-accounts/australian-national-accounts-national-income-expenditure-and-product/jun-2024
8. COVID-19 Mortality in Australia: Deaths registered until 31 January 2024 | Australian Bureau of Statistics [Internet]. 2024 [cited 2024 Sep 29]. Available from: https://www.abs.gov.au/articles/COVID-19-mortality-australia-deaths-registered-until-31-january-2024
9. Australian Institute of Health and Welfare [Internet]. 2024 [cited 2024 Sep 29]. COVID-19. Available from: https://www.aihw.gov.au/reports/australias-health/COVID-19
10. COVID-19 Vaccine Rollout [Internet]. Department of Health; 2021 [cited 2024 Sep 29]. Available from: https://www.health.gov.au/sites/default/files/documents/2021/10/COVID-19-vaccine-rollout-update-20-october-2021.pdf
11. CDGN [Internet]. [cited 2024 Nov 3]. Variants of concern. Available from: https://www.cdgn.org.au/variants-of-concern

Supplementary R code to produce visualisations

```{r}

library(magrittr)
library(tidyverse)
library(plotly)
library(colorspace)
library(data.table)
library(patchwork)

```

```{r}

# Load in data
# COVID 19 dataset was retrieved from
https://ourworldindata.org/coronavirus/country/australia
covid <-
  fread("covid.csv",
        header = TRUE) %>%
  tibble() %>%
  filter(location == "Australia")

# Economic data
# Retrieved from
https://www.abs.gov.au/articles/COVID-19-mortality-australia-deaths-registered-until-31-january-2024
gdp <-
  fread("economy_gdp_change.csv",
        header = TRUE) %>%
  tibble() %>%
  mutate(date = my(date) + month(1),
         gdp_change = gdp_change / 100)

```

```{r}
# Data inspection
names(covid) %>% grep(pattern = "new", value = TRUE)

# Data cleanup
```

```
# Group data by yearly quarter
covid_aus <-
  covid %>%
  mutate(date = ymd(date)) %>%
  group_by(date = floor_date(date, "month")) %>%
  summarise(
    ## total_cases = sum(total_cases),
    ## total_deaths = sum(total_deaths),
    ## total_vaccinations = sum(people_fully_vaccinated),
    new_cases = sum(new_cases_per_million),
    ## new_deaths = sum(new_deaths),
    new_vaccinations = sum(new_vaccinations_smoothed_per_million),
  ) %>%
  replace(is.na(.), 0) %>%
  pivot_longer(!date,
               names_to = "stat")

covid_aus_cum <-
  covid %>%
  mutate(date = ymd(date)) %>%
  group_by(date = floor_date(date, "3 month")) %>%
  summarise(
    total_cases = sum(new_cases) %>%
      as.integer(),
    total_hospitalised = sum(hosp_patients) %>%
      as.integer(),
    ## total_deaths = sum(new_deaths_smoothed) %>%
    ##   as.integer(),
    total_vaccinated = sum(new_people_vaccinated_smoothed) %>%
      as.integer(),
  ) %>%
  replace(is.na(.), 0) %>%
  mutate(total_cases = cumsum(total_cases),
         total_hospitalised = cumsum(total_hospitalised),
         ##
         total_vaccinated = cumsum(total_vaccinated),
         ) %>%
  pivot_longer(!date,
               names_to = "stat")

covid_aus$stat <-
  factor(covid_aus$stat,
         labels = covid_aus$stat %>%
           unique() %>%
```

```
          gsub("[_]", " ", .) %>%
          str_to_sentence())

covid_aus_cum$stat <-
  factor(covid_aus_cum$stat,
         labels = covid_aus_cum$stat %>%
           unique() %>%
           gsub("[_]", " ", .) %>%
           str_to_sentence())

```

```{r}

# Number of cases in Australia
p_cases <-
  ggplot(data = covid_aus %>%
           filter(date >= ymd("2021-01-01")),
         aes(x = date,
             y = value,
             fill = stat)) +
  geom_col(width = 25.0) +
  scale_x_date(date_breaks = "3 month",
               expand = c(0.03, 0),
               date_labels = "%Y-%m") +
  scale_y_continuous(expand = c(0, 0),
                     labels = scales::number_format(),
                     ## sec.axis = sec_axis(),
                     ) +
  theme_classic() +
  theme(legend.position = "inside",
        legend.position.inside = c(0.5, 0.95)) +
  guides(fill = guide_legend(NULL,
                             direction = "horizontal")) +
  xlab(NULL) +
  ylab("Number of vaccinations/cases per million") +
  ggtitle("COVID 19 impact on Australia")

p_economy <-
  ggplot(data = gdp %>%
           filter(date >= ymd("2021-01-01")),
         mapping = aes(x = date,
                       y = gdp_change,
                       fill = gdp_change)) +
```

```
  geom_col(width = 75.0) +
  scale_x_date(date_breaks = "3 month",
               expand = c(0.03, 0),
               date_labels = "%Y-%m") +
  scale_y_continuous(expand = c(0, 0),
                     labels = scales::percent_format()) +
  scale_fill_continuous_diverging(palette = "berlin",
                                  rev = TRUE,
                                  labels = scales::percent_format()) +
  theme_classic() +
  theme(axis.title.x = element_blank(),
        axis.text.x = element_blank(),
        axis.ticks.x = element_blank(),
        axis.line.x = element_blank(),
        legend.position = "inside",
        legend.position.inside = c(0.5, 0.95),
        legend.key.width = unit(0.05, "npc"),
        ) +
  guides(fill = guide_colorbar(NULL,
                               direction = "horizontal")) +
  xlab(NULL) +
  ylab("Quarterly GDP growth")

p_combined <-
  wrap_plots(p_cases,
             p_economy,
             ncol = 1,
             heights = c(3, 2)) +
  scale_x_date(date_breaks = "3 month",
               expand = c(0.03, 0),
               date_labels = "%Y-%m") +
  scale_y_continuous(expand = c(0, 0),
                     labels = scales::percent_format()) +
  scale_fill_continuous_diverging(palette = "berlin",
                                  rev = TRUE,
                                  labels = scales::percent_format()) +
  theme_classic() +
  theme(axis.title.x = element_blank(),
        axis.text.x = element_blank(),
        axis.ticks.x = element_blank(),
        axis.line.x = element_blank(),
        legend.position = "inside",
        legend.position.inside = c(0.5, 0.95),
        legend.key.width = unit(0.05, "npc"),
```

```
      ) +
  guides(fill = guide_colorbar(NULL,
                               direction = "horizontal")) +
  xlab(NULL) +
  ylab("Quarterly GDP growth")

ggsave("cases_economy.png",
       plot = p_combined,
       width = 12,
       height = 8)

```

```{r}

p_cumulative <-
  ggplot(data = covid_aus_cum,
         mapping = aes(x = date,
                       y = value / max(value),
                       fill = stat)) +
  geom_col(position = "dodge",
           stroke = 1.0,
           width = 50.0,
           alpha = 0.5) +
  geom_smooth(mapping = aes(color = stat),
              formula = y ~ x,
              method = "glm",
              method.args = list(family = "quasibinomial"),
              se = FALSE,
              show_guides = FALSE) +
  scale_x_date(date_breaks = "3 month",
               expand = c(0.03, 0),
               date_labels = "%Y-%m") +
  scale_y_continuous(expand = c(0, 0),
                     n.breaks = 5,
                     labels = scales::number_format(accuracy = 1e6,
                                                    scale = max(covid_aus_cum$value)),
                     ## transform = scales::transform_log10(),
                     ) +
  scale_fill_brewer(type = "qual", palette = 1) +
  scale_color_brewer(type = "qual", palette = 1) +
  theme_classic() +
  theme(legend.position = "inside",
        legend.position.inside = c(0.10, 0.90),
```

```
        legend.key.width = unit(0.05, "npc"),
        ) +
  guides(fill = guide_legend(NULL,
                             direction = "vertical"),
         color = guide_legend(NULL)) +
  xlab("Date") +
  ylab("Population cases") +
  ggtitle("Combined COVID 19 healthcare impacts within Australia")

ggsave("cases_cumulative.png",
       plot = p_cumulative,
       width = 12,
       height = 8)
```