

Vaccination Rate Influencing Victoria

Group 63 - COMP20008

Zihan Xu , Yingyi Luan , Yuntao Lu , Xueying Sun

Frame the problem

Question

The question is when is the time to lift restrictions. Collecting data from the **health system of Victoria is needed to be the first step**. We can **predict the appropriate time** when it is safe for people to go wherever they want without restrictions by evaluating the number of **confirmed cases, fully vaccinated people and active cases**. In other words, the infectious people(especially those severely infected) can be treated properly with the spread of vaccination and enhancement of medical systems. The appropriate point to lift restrictions is closely connected with the health of citizens, and it can guarantee the recovery of the economy without bearing risk of safety.

Dataset

Name	Content	What we extract
'COVID_AU_state_cumulative till 10.05.csv'	State, confirmed, death, tests, positives, recoveries, hospitalisation, icu, vent and vaccines	Confirmed cases and vaccines in VIC and NSW
'COVID_AU_state_daily_change.csv'	Same as above	Same as above
'2 Doses.csv'	Fully vaccinated number, sorted by States	VIC and NSW data
'Active cases.csv'	Active cases in Victoria with dates.	All
"Cases in ICU.csv"	ICU cases for each state with dates.	VIC and NSW data
"Active cases in NSW.csv"	Active cases in New South Wales (NSW) with dates	All

nurse.csv	Health employee in countries	Standardised data from 2017 to 2019
health_stat.csv	Health status in countries	Same as above
hospital.csv	Num of hospital in countries	Same as above
hospital_aggr.csv	Hospital aggregation in countries	Same as above
life_expentancy.csv	Life expectancy in countries	Same as above
meditech.csv	Medical technology in countries	Same as above
recovery.csv	Japan's covid-19 data	Japan's required inpatient cases

Data Wrangling and Analysis:

Victoria

Figure 1 includes scatter plots of the cumulative number of confirmed cases and the number of fully vaccinated people from 15th September to 3rd October. We chose this period of time because it is more recent, within this period, both the confirmed cases and fully vaccinated number increase.

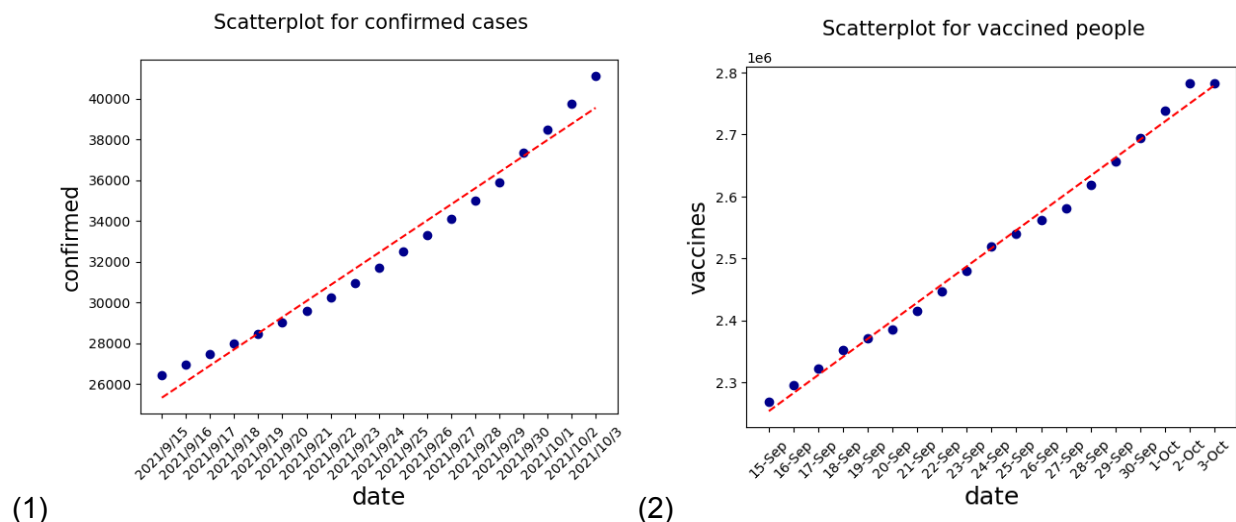
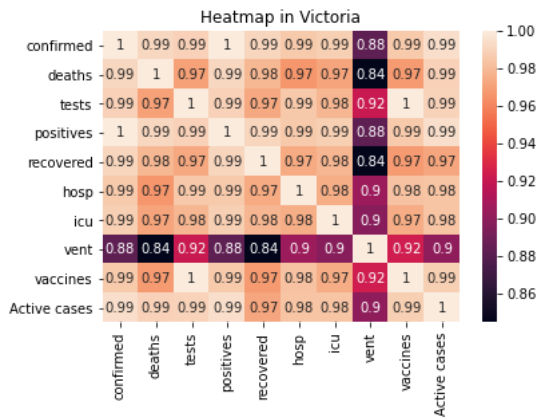
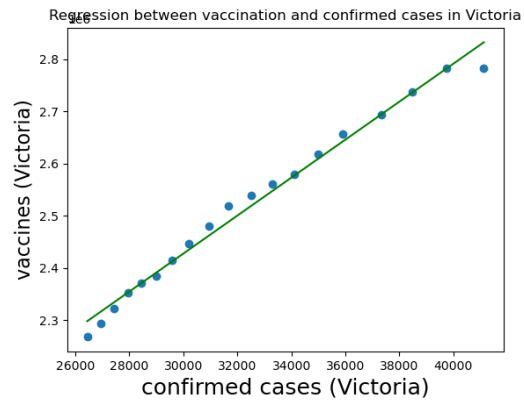


Figure 1: (1)scatterplot of confirmed cases in Victoria with fitted line $y = 790.1x + 2.532e+04$. (2)scatterplot of fully vaccinated numbers in Victoria with fitted line $y = 2.922e+04x + 2.253e+06$. All data here is not logged.



(1)



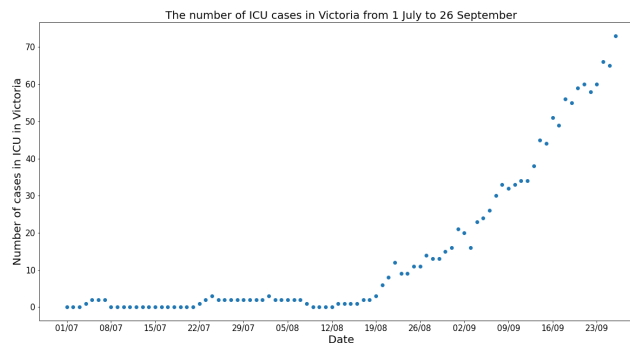
(2)

Figure 2: (1) heatmap for data collected in Victoria, x-axis (confirmed cases) as observed value and y-axis (2-dose vaccines) to be predicted value. (2) Scatterplot between fully vaccinated people and confirmed cases in Victoria and a regression line.

Implications

- Elements are closely related and influenced.
- ICU cases go up as confirmed cases grow.
- Confirmed cases are the indicator for citizens to evaluate the seriousness of epidemics and the risk of getting infected. As confirmed cases increase, citizens may decide to get vaccinated due to their personal self-protection.
- The reason why the correlation between vaccinations and active cases is negative is probably because people are scared about the growth of cases and are encouraged to take vaccines

Figure 3 indicates the number of ICU cases in Victoria. There are only a few ICU cases from July 1 to August 19 in the first scatterplot. This data is not sufficient for analysing and making predictions. Therefore, we rearrange the period to August 15 to September 26. The second scatterplot shows a positive, linear relationship between dates and Number of cases in ICU in Victoria. There is **no obvious outlier** in this dataset.



1)

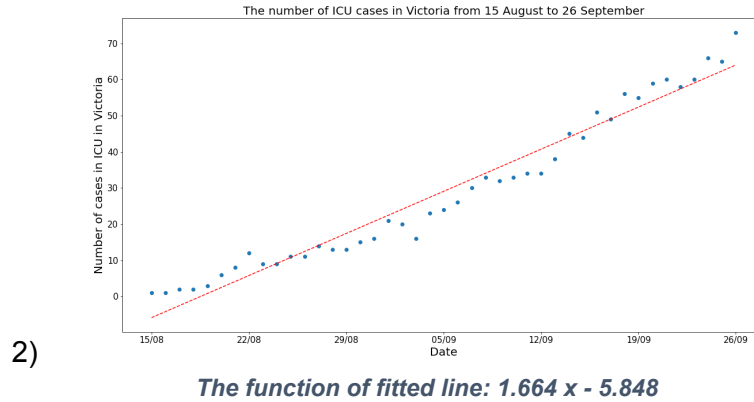


Figure 3: (1) Scatterplot of the number of active cases in Victoria. (2) Scatterplot of the number of active cases in Victoria with a fitted line.

Figure 4 illustrates the number of active cases in Victoria. Originally, the left scatterplot shows the number of active cases grows exponentially. To detect the outlier, linearize the graph on the left, then it comes to the result as right of the graph and there is **no obvious outlier as shown**.

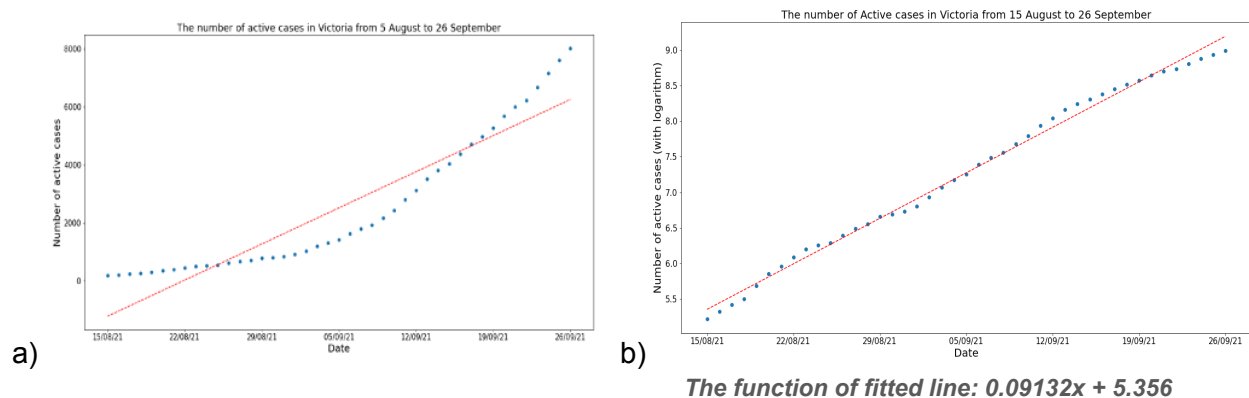


Figure 4: (a) Scatterplot of the number of active cases in Victoria. (b) Scatterplot of the number of active cases with logarithm.

It can be seen that data of Victoria are biased and not sufficient to analyse, because it might be interrupted by externalities such as protesting events. These can be shown by the fact that vaccination cases go up the same way as active cases, which seems like the vaccination VIC residents take are not enough to give them protection, we have to use other states' statistics. Therefore, we also extract confirmed cases in NSW for comparison, reference and prediction from 1st July to 30th September 2021 from COVID_AU_state_cumulative till 10.05.csv. Also as the ICU cases in Victoria in July, there is no reference value for analysing.

New South Wales (NSW)

Figure 5 includes scatter plots illustrating the trend of cases confirmed in July, August, September of NSW, both of the July and August one are **exponentially increasing** while in September, it can be clearly seen that the trend tends to be **linear with slightly decreasing**

gradient from 25th September 2021. This figure also includes scatter plots indicating the trend of two-dose vaccines given in July, August and September of NSW, the vaccination trend keeps linearly increasing in these three months.

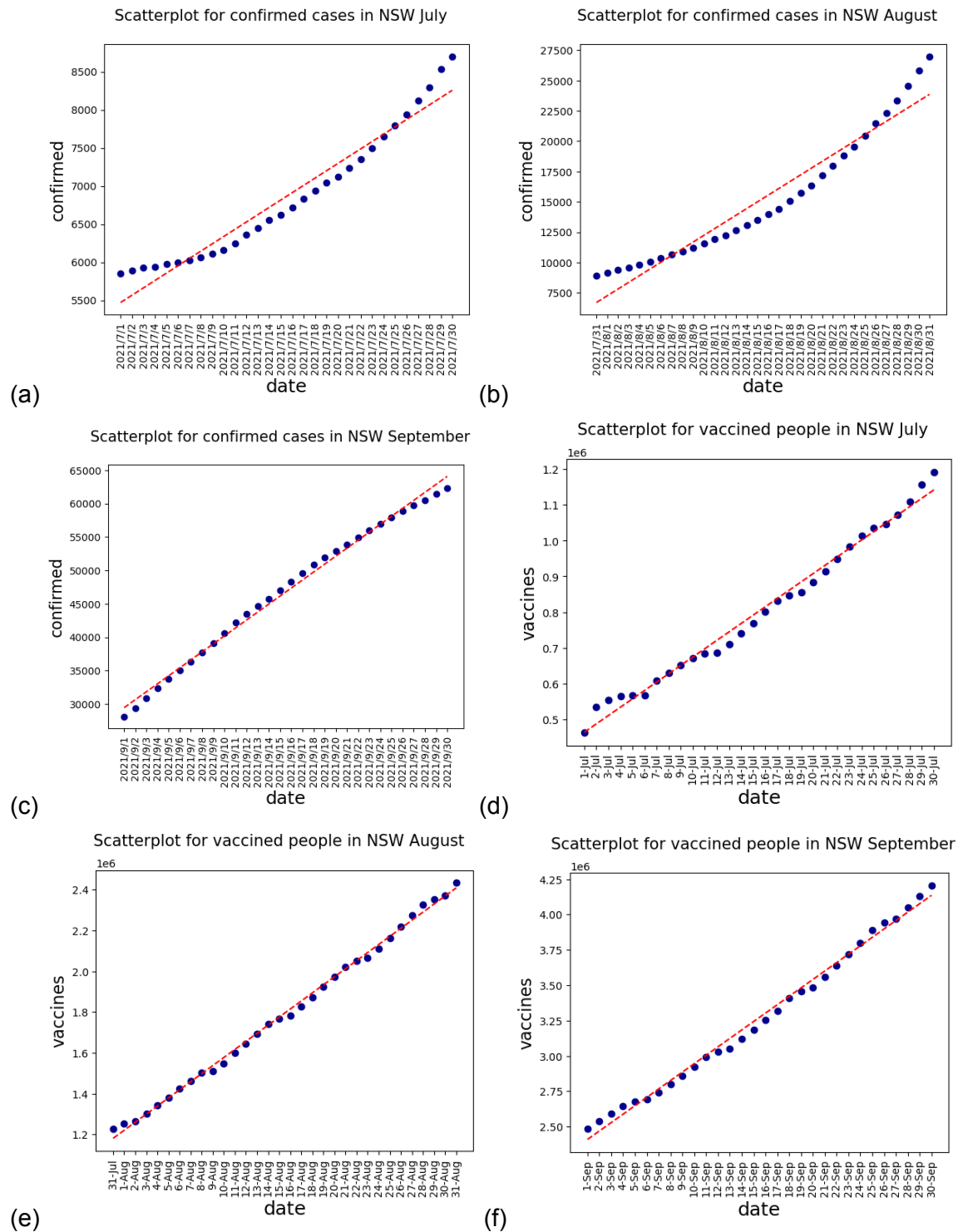


Figure 5: the number of confirmed cases and number of fully vaccinated people in July, August and September in NSW. (a) $y = 96.05x + 5472$ (b) $y = 554x + 6694$ (c) $y = 1195x + 2.943e+04$ (d) $y = 2.334e+04x + 4.644e+05$ (e) $y = 3.962e+04x + 1.181e+06$ (f) $y = 5.956e+04x + 2.409e+06$

Figure 6 shows the relationship between dates and number of ICU cases in NSW. As the figure showed, up to September 15, the number of ICU cases increased linearly. Then, on a downward trend since 22 September.

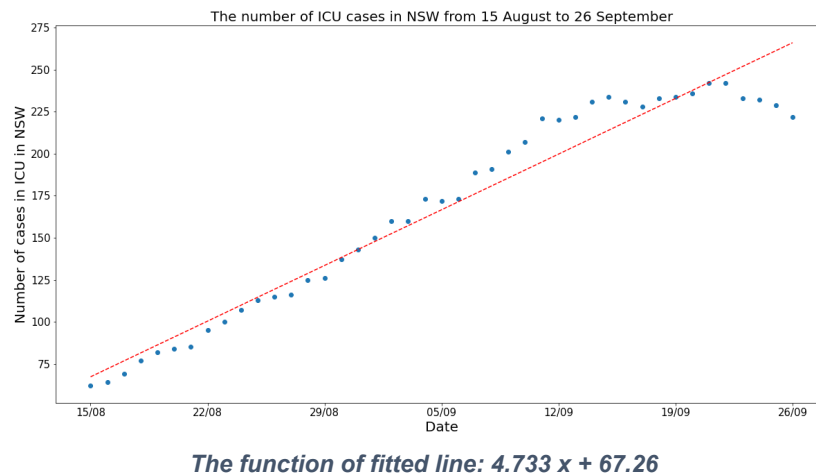
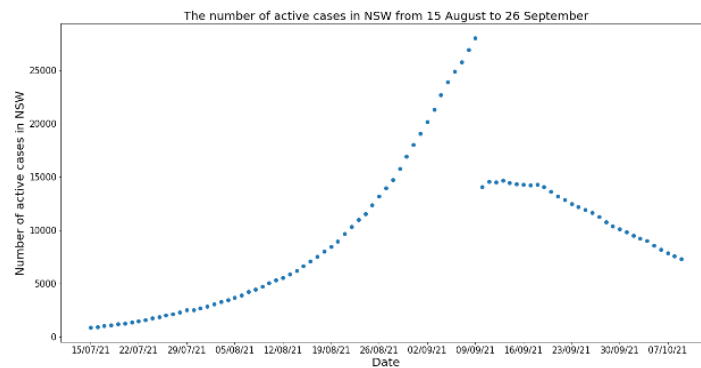


Figure 6: Scatterplot of the number of ICU cases in NSW with a fitted line.



a)

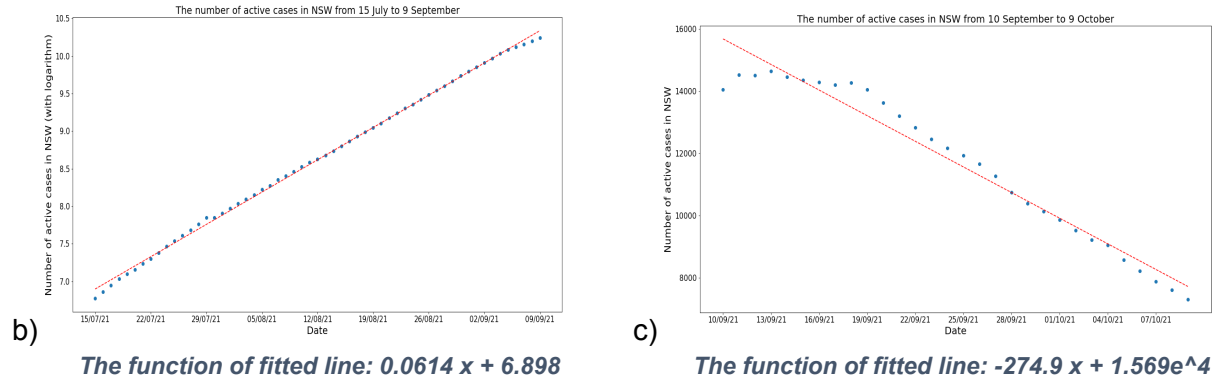


Figure 7: (a) Scatterplot of the number of active cases in NSW. (b) The number of active cases in NSW from July 15 to September 9 with logarithm. (c) The number of active cases in NSW from 10 September to 9 October.

Perform in-depth analysis

According to last step, we get the following discoveries

- Both fully vaccination people in Victoria and active cases goes up **exponentially**
 - Note: Active case = accumulative confirmed case - total recovery - death
- The daily confirmed cases in Victoria goes up **linearly**
 - This may implicate the Victorians are scared by the growth of confirmed cases and are encouraged to take vaccinations
- The number of ICU cases goes up **linearly**
 - It indicates that the active cases and confirmed cases can result in the growth of ICU cases

In this case, we can not get the answers only through researching Victoria data. According to common sense, vaccination can to some extent protect us from covid. The daily confirmed cases and active cases can goes down by that time

Since NSW is very similar to VIC in various areas and NSW goes earlier than VIC in vaccination and confirmed cases, we are going to do research on NSW data and find the predictive models, and use that model to predict VIC data.

By analysing data in NSW, we can get the following conclusions

- The active cases in NSW goes up **exponentially** with fully vaccinated population at first and then goes down **linearly** (after 9 Sep)
 - There is a large jump just before it goes down, which means there may exists a huge error for data
 - There are many possible reasons contributing to its decrease, eg. large number of recovery cases/ confirmed cases goes down dramatically, which makes it hard to imply

The decrease of active cases leads the reduction in the number of ICU cases

- The number of ICU cases reaches **a summit** on approximately 20 September, and goes down afterward
 - The peak of ICU cases may comes after daily active cases/confirmed cases' peak

In order to do a regression and make predictions, we have to explore the correlation first. The heatmap shows a relatively high correlation for vaccines and confirmed cases with 0.98, and the correlation coefficient between active cases and ICU cases is 0.75, which still is a highly positive correlation.

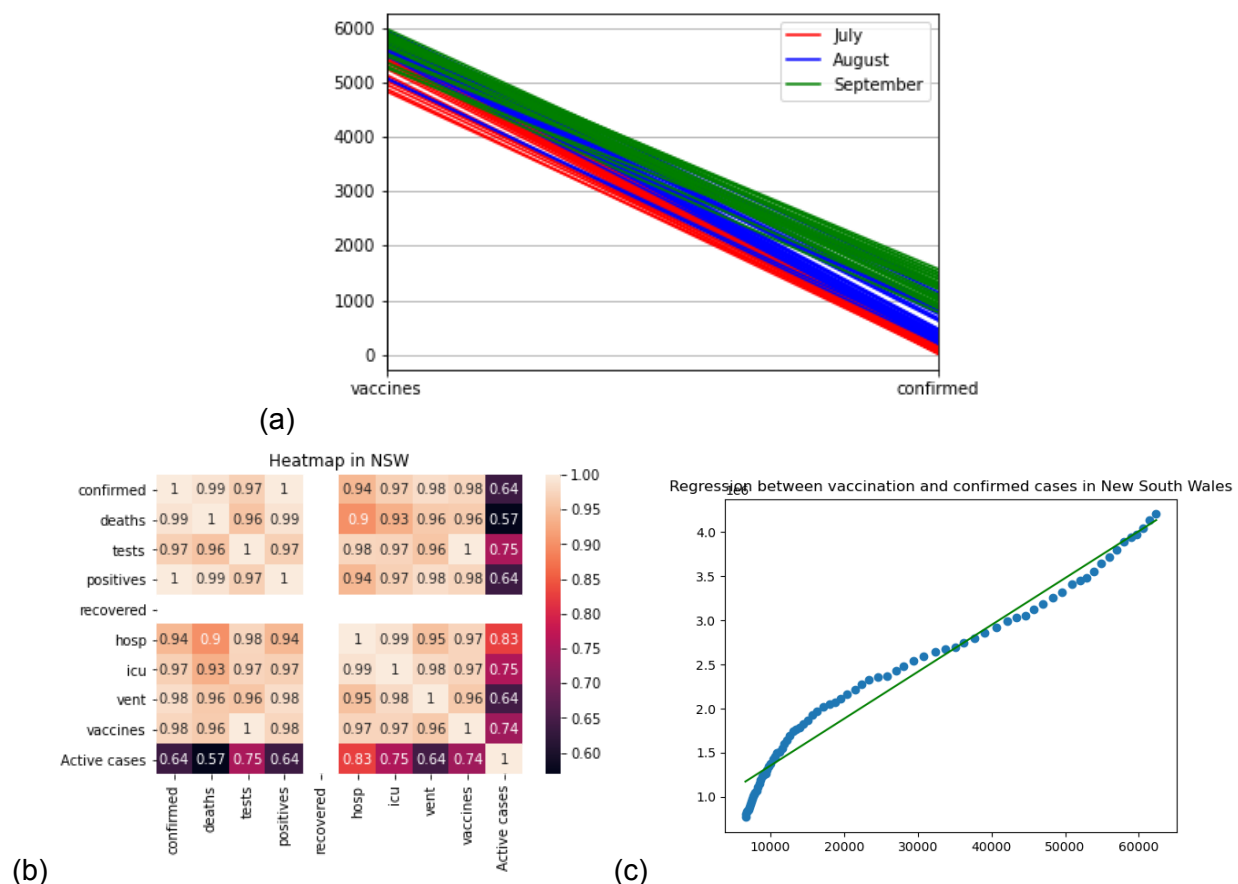


Figure 8: (a)parallel plot for daily changed confirmed cases and vaccines (b)the heatmap of correlation among different elements of recorded data during the period of COVID-19 and (c)the scatterplot between two-dose vaccines given and confirmed cases with a regression line.

Prediction Models

The prediction model we build is to use the **fully vaccinated population**(after 15/7/2021) to predict the **daily confirmed cases**(after 9/9/2021), and then to use the daily confirmed cases predicted to predict the icu cases, and test whether the icu cases is within the Victoria capacity

- **Fully vaccinated population VS daily confirmed cases**

- The regression model we get is

Daily confirmed cases = $-214.4028283 * (\text{days passed}) + 9407.94226747 * \log(\text{vaccinated population}) - 127317.28683888628$

- The Coefficient of determination for training data is 0.91, and for test is 0.79, which means this model is very persuasive

- **Daily confirmed cases VS ICU cases**

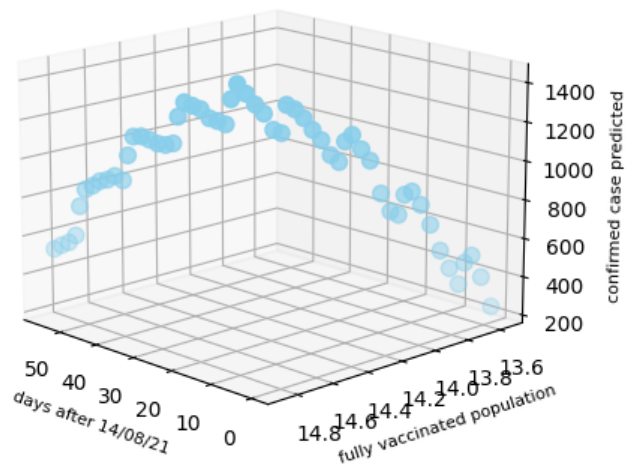
- a. The regression model we get is

ICU cases = $3.23409702 * \text{days passed} + 0.05498894 * \text{daily confirmed cases} + 19.25122357429609$

- b. Coefficient of determination (test): 0.90
Coefficient of determination (training): 0.91

This model is well fitted either in data side(r squared through graph)

he regression between confirmed cases and fully vaccinated population in NSW from 10/8/2



The regression between icu cases and confirmed cases in NSW from 10/8/2021

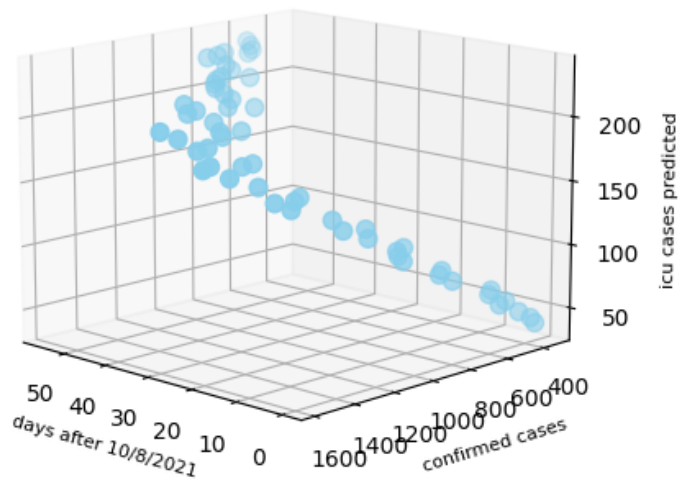


Figure 9: Two regression models.

- **OUR prediction**

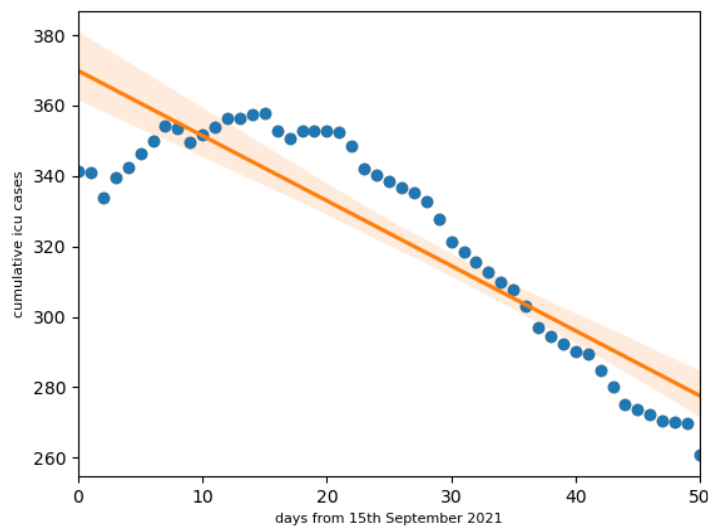
According to our prediction model, Victorian's daily confirmed cases will reach 5000+ before going down stably, and icu's cumulative cases will reach 360 before going down, accumulative ICU cases's peak will come about 20 days after confirmed cases' peak

- **Possible problems**

There exist some errors of our prediction according to current data

- Victorian's situation may be different from NSW's in various ways, Which means some error will appear if we directly apply NSW's model In prediction
- We only have less than 3 months' valid data for this model, which may produces error for the model itself

The scatterplot of predicted icu cases in Victoria from 15/9/2021



The scatterplot of predicted confirmed cases in Victoria from 15/9/2021

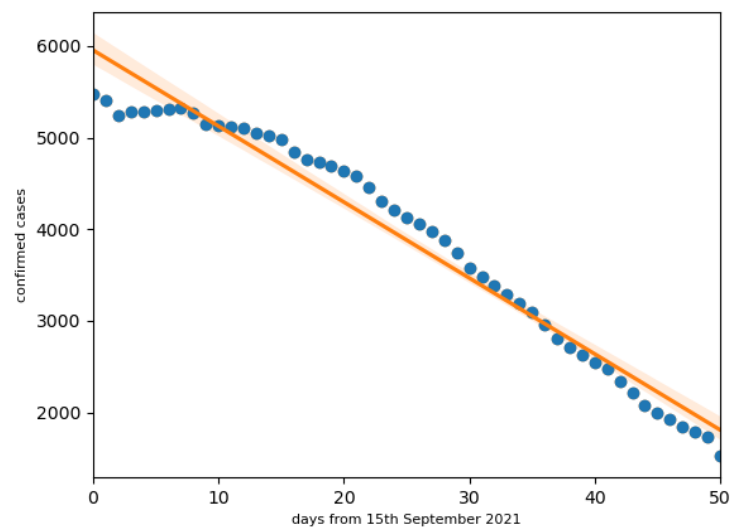
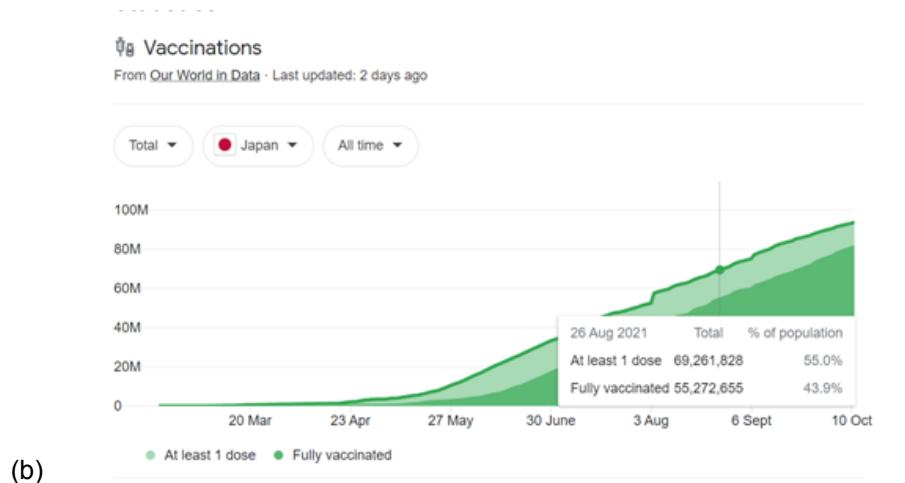
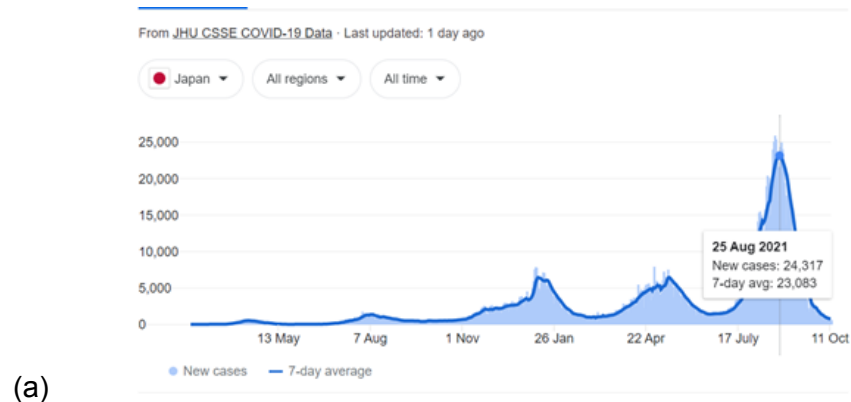
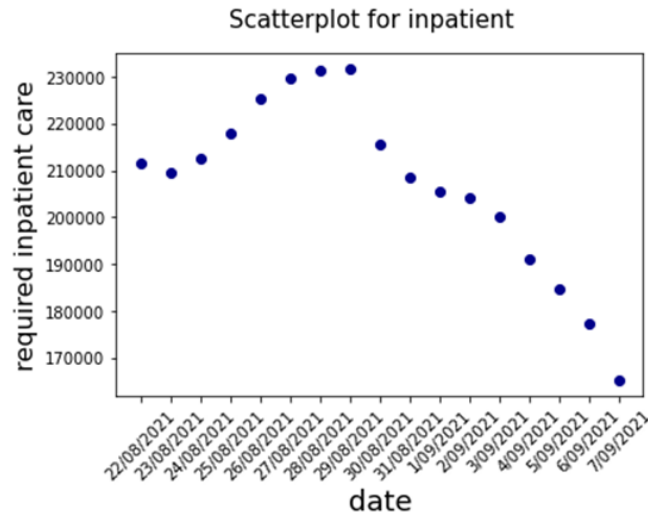


Figure 10: Two prediction models.

Further Research

To make the data analysis more complete, based on the k-means clustering on health indicator dataframes, we find Japan's health system is similar to Australia. Therefore we find Japan's covid-19 related data compared with NSW and VIC. From figure 8, we know the trend of covid-19 in Japan. The required inpatient cares peak 3 days after the daily new cases peaked. Around the peak, 44% of the population had been fully vaccinated, and 55% had received their first dose.





(c)

Figure 11: (a). new cases daily in Japan (from google) (b). Vaccination rate and numbers in Japan (from google) (c). cumulative required inpatient cares in Japan.

Result and significance:

By analysing data above, we are confident to say that the confirmed and active cases in Victoria will go down in the near future with such a large proportion of people vaccinated (60+% for fully vaccinated), and the ICU cases will decrease as well.

According to official statistics

NSW completed the 70% first dose of COVID-19 vaccination in the eligible population on 2 September 2021

By 5 September 2021, 40% of the NSW population was fully vaccinated

The huge gap between first and second dose in NSW also shows that many citizens get their first dose in August, yet haven't got their second dose in early September.

By referring to Japan's data and combining it with the data in NSW, we notice at least 40% of residents were fully vaccinated, the peak will occur and the burden of the health system will be reduced, since the required inpatient care cases are reduced.

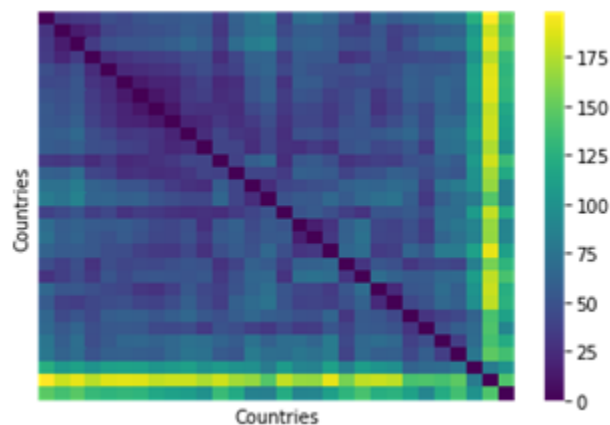
Conclusion:

The number of ICU beds is 1130, our prediction is based on the vaccination rate producing the result that confirmed cases peak is 5000 and ICU peak is 300. Therefore, Victoria's health system is enough to support severe patients. However, to avoid insufficient data, we also

referred to NSW and Japan's dataset. We found that when vaccination rate reaches 40%, the active cases will go down. This result revealed that Under the condition of active vaccination, when will the Covid-19 in Victoria get better. This will help the government to make better decisions to promulgate sensible policies to help the economy recover better.

Limitations:

1. When we research on dataset or analyse the result, we chose to ignore how it may change in different age group or sex group, we assigned them as control.
2. When we analysed the similarity of health system, we also plot heatmap of these countries shown below:



This figure shows the clusters might be inaccurate due to the insufficient data we use. Besides, although the health system tends to be similar with Japan, the policy is still different towards covid-19. Both Victoria and NSW never had over 10000 daily new cases, but Japan did. Therefore, it's hard to tell the importance of vaccination. The reduction might be caused by herd immunity in Japan.

Reference:

https://en.wikipedia.org/wiki/COVID-19_vaccination_in_Australia

Source of dataset:

1. COVID_AU_state_cumulative till 10.05.csv and COVID_AU_state_daily_change:
https://github.com/M3IT/COVID-19_Data/blob/master/Data/COVID_AU_national_daily_change.csv
2. 2 Doses, Active cases, Active cases in NSW and Cases in ICU:
<https://www.covid19data.com.au/victoria>
3. Health quality indicators (health status, hospital, health employee, hospital aggregations (Utilization), life expectancy and medical technology)
[Health Care Quality Indicators \(oecd.org\)](https://www.oecd.org/health/health-care-quality-indicators/)
4. Inpatient care cases in Japan
<https://covid19.mhlw.go.jp/extensions/public/index.html>
5. New cases and vaccination rate in japan
https://www.google.com/search?q=japan+covid+19+&sxsrf=AOaemvleyX7caFQZjAlZvolKrylyxTFxrA%3A1634256301163&ei=rcVoYYm4CYaY4-EP556HsAo&ved=0ahUKEwjJwrWSj8vzAhUGzDgGHWfPAaYQ4dUDCA4&uact=5&oq=japan+covid+19+&gs_lcp=Cgdn d3Mtd2l6EAMyBAgjECcyBAgjECcyCAgAEMkDEJECMgUIABCRAjIHCAAQsQMqqzIECAAQQzIFCAAQgAQyBQgAEIAEMgUIABCABDIFCAAQgAQ6BwgjELADECC6BwgAEEcQsANKBAhBGABQIRIYIRlg2xtoAXACeACAAaMBiAGjAZIBAzAuMZgBAKABAcgBCcABAQ&scient=gws-wiz
6. ICU beds in victoria
<https://www.parliament.vic.gov.au/assembly/questions-on-notice/questions-database/details/53/3582>