THE UNIVERSITY OF SYDNEY

Assignment Coversheet ☐ GROUP ASSIGNMENT

Please fill in your details below. Use one form for each group assignment.

Assignment Details:

Assignment Title	**COMP5048 GROUP REPORT**				
Assignment number	**2**				
Unit of Study Tutor	**Anjali Singh**				
Group or Tutorial ID	**RETUT01-GROUP 4**				
Due Date	**NOV, 05, 2021**	Submission Date	**NOV, 04, 2021**	Word Count	2985

Membership

Group Name/Number	**RETUT01-GROUP 4**				
Family Name	Given Name (s)	Student Number (SID)	Unikey	Contribution + percentage	Signature
Wenzhao	Liu	510069652	wliu6317	20%	Wenzhao Liu
Jingyi	Wu	510222730	jiwu4151	20%	Jingyi Wu
Qinqin	Zhu	510170590	qzhu3515	20%	Qinqin Zhu
Chenzhen	Li	510233286	chli7330	20%	Chenzhen Li
Yichen	Chen	510138903	yche3494	20%	Yichen Chen

Declaration:

1. I understand that all forms of plagiarism and unauthorised collusion are regarded as academic dishonesty by the university, resulting in penalties including failure of the unit of study and possible disciplinary action.
2. I have completed the **Academic Honesty Education Module** on Canvas.
3. I understand that failure to comply with the Academic Dishonesty and Plagiarism in Coursework Policy can lead to the University commencing proceedings against me for potential student misconduct under Chapter 8 of the ***University of Sydney By-Law 1999*** (as amended).
4. This work is substantially my own, and to the extent that any part of this work is not my own I have indicated that it is not my own by acknowledging the source of that part or those parts of the work.
5. The assessment has not been submitted previously for assessment in this or any other unit, or another institution.
6. I acknowledge that the assessor of this assignment may, for the purpose of assessing this assignment may:
   a. Reproduce this assignment and provide a copy to another member of the school; and/or
   b. Use similarity detection software (which may then retain a copy of the assignment on its database for the purpose of future plagiarism checking).
7. I have retained a duplicate copy of the assignment.

Please type in your group number here to acknowledge this declaration:	**RETUT01-GROUP 4**

COMP5048 Group Report (RETUT01 – G04)

Jingyi Wu
COMP5048
The University of Sydney
Sydney, Australia
SID: 510222730

Yichen Chen
COMP5048
The University of Sydney
Sydney, Australia
SID: 510138903

Wenzhao Liu
COMP5048
The University of Sydney
Sydney, Australia
SID: 510069652

Qinqin Zhu
COMP5048
The University of Sydney
Sydney, Australia
SID: 510170590

Chenzhen Li
COMP5048
The University of Sydney
Sydney, Australia
SID: 510233286

PART A: GOOD HYGIENIC PROCESSES

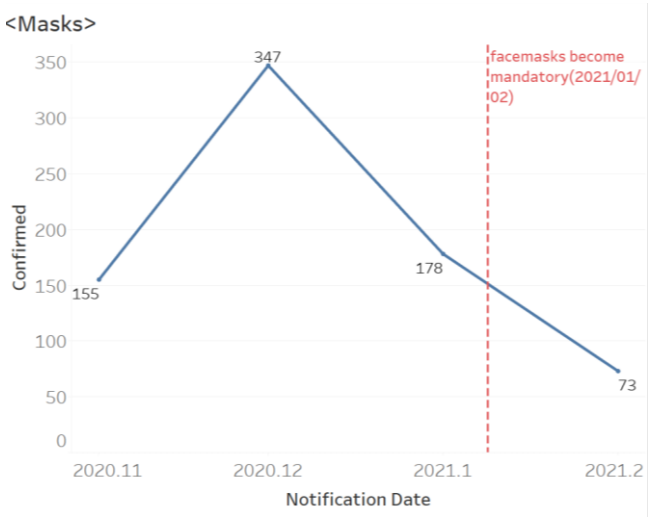

Figure 1

In this part, the line chart is selected for visual analysis of the COVID-19 epidemic, because the line chart can clearly show the trend of the number of confirmed cases over time, which is simple, direct and easy to observe. The horizontal axis is time, and the vertical axis is the number of confirmed cases.

As is shown in Figure 1, mandatory mask wearing on January 2, 2021 led to a downward trend in COVID-19 diagnoses over the following month. So good hygiene procedures (wearing masks) can have a positive impact on COVID-19 prevention and control the spread, reducing the number of people diagnosed with COVID-19.

PART A: ENCOURAGING TEST:

We know that since the beginning of the epidemic, COVID-19 testing has been encouraged.

To determine the extent of "encouragement" more clearly, we created a line chart of test numbers versus dates (Figure 2) to find out if there is/are obvious dates when the New South Wales government has significantly increased the testing efforts.

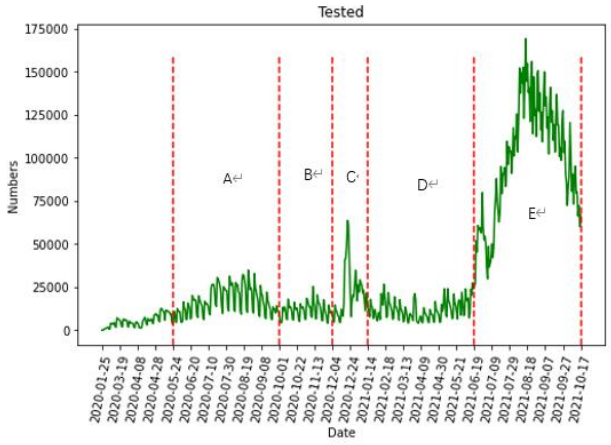

Figure 2

For area A, B, C, D and E, we can see that the A,C, and E are obviously mountain-shaped. In these three periods, the number of tests are obviously greater than that of B and D, so we think that the time periods of A, C and E sections' testing were highly encouraged by the government. Large detection effort. We select the C section and determine whether the confirmed number of next month (that is, from January 15 to February 15, 2021) will be lower than the previous month (From December 15th to 15th January 2021) .

The number of tests is ratio data, because the number of tests can be true 0, and the date is ordered data. They are homogeneous. We represent the date via x -axis and using y-axis to represent the number of tests. We can observe the trend in a period of the line chart to easily determine whether in this period the number of tests increased and when it started to decrease.

We use the line chart because it can clearly reflect whether the number of people inspected becomes more or less over time. Compared with the bar chart, the line chart has continuous characteristics, which can help us find multiple time periods which encourage testing and when does the time period begin and when it ends.

We build a bar chart Figure 3 based on the number of confirmed cases versus date. Set a red dotted line in where x = 2021-01-15,

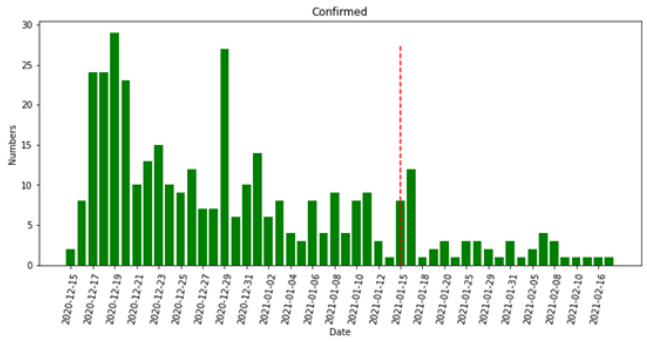

Figure 3

We can clearly see that if we regard the red dotted line as the boundary, most of the columns on the right have values less than 5, while the columns on the left have larger values, which are generally greater than 5, and even more than 20 on many days. Through the visualization of the bar chart, we can determine that the number of confirmed cases in the month after January 15 is much less than that in the month before January 15, which means that it is effective to encourage testing.

We use a bar chart because it can more clearly show the size of a certain point in date, and as we only focus on whether the number of confirmed numbers per day of the month before and after January 15 is larger or smaller in total, we do not care about how the numbers changed by time like it is increase first or decrease first, so we did not use a line chart.

Confirmed numbers are ratio data, because it can be true zero, and date is internal data, they are homogeneous. We used x axis to represent date, y axis to represent numbers of confirmed, determine whether the numbers of testing are high or low via the height of columns in different sides of red dotted line.

As mentioned above, we found several time periods that test numbers significantly increased via a line chart and found that encouraging testing can reduce the number of confirmed via a bar chart. We expect that this measure (or combination with other measures at the same time period) can help us get rid of COVID19. But it seems that COVID 19 re-spread after June 28, 2021, after the delta virus raged. We will use a line chart Figure 10 (PS: Figure 9 is shown following report, this is a comprehensive chart of three possible ineffective measures) to confirm this hypothesis, using the number of confirmed diagnoses from January 2020 to October 2021. We have inserted a yellow dotted line after the C time period (2021-01-15) we just used as an example above. We can clearly see that after implementing measures to encourage testing, the number of confirmed cases has stabilized at a low level, but after June 28, 2021, there has been an explosive increase in the number of confirmed cases, which is higher than any historical time. This is ample evidence that the series of measures before that date, including encouraging testing, did not completely eliminate the epidemic as we expected. We are not saying that they are useless, but with the emergence of the delta virus, they can no longer play an effective role as we expected, which means that we may need stronger measures.

PART A : TRAVEL RESTRICTIONS — INTERSTATE BORDER:

For exploring whether the controlling flow of people in NSW has an effect on numbers of infected people, we picked two key points. One was on 25 June that the NSW had implemented a lockdown with stay-at-home restrictions, the other one was curfew on 23 August. They were represented as blue dashed lines separately in the visualization, and were marked with their date next to the dashed lines. In order to judge the effect of this approach, we decided to use line graphs in visualization, since we would like to display how the number of infected people change over this period. For designing the visualization, the x-axis is the timeline which is the period between July and October in 2021, and the y-axis is the number of infections. Then we would like to use line graphs here. Since, the line graph could showcase the trend. If the trend is downward, it indicates that this measure has a good effect on controlling the COVID-19.

As we can obtain from this visualization Figure 4, at the beginning of the implementation of this restriction, the control of infections is not obvious, the trend of the number of infections is still rising. However, we consider that there is a lag in the effect of the lockdown. After two weeks, the number of infected people here has still not decreased significantly which indicates that the effect of the first implementation of this lockdown restriction is not very good. After the curfew was implemented as well in August, there was a turning point in September within two weeks of the implementation of this measure. The number of infections began to decline gradually, which indicates that the restriction in NSW would have an effect in controlling the Delta, although there is a lag in its effects.

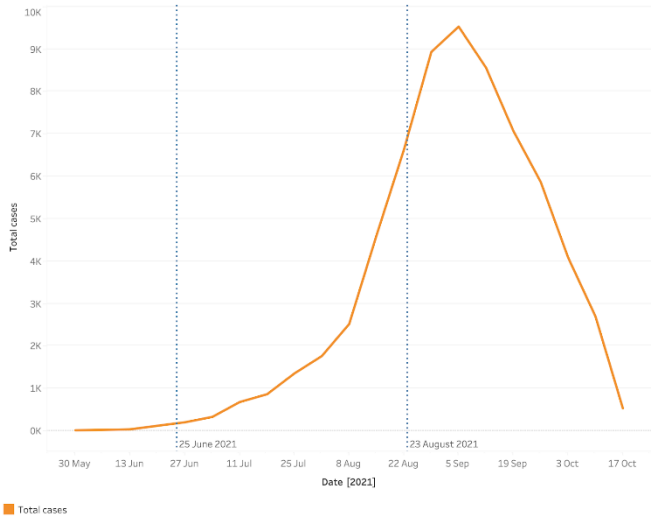

Figure 4

PART A : TRAVEL RESTRICTIONS — INTERNATIONAL BORDER:

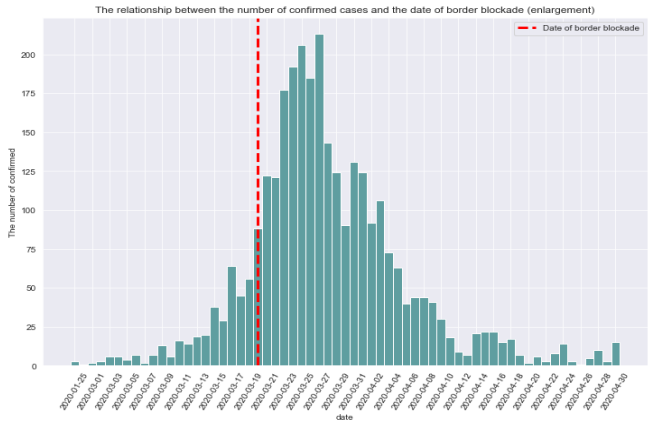

Figure 5

For the strategy of controlling the flow of people through international border control, we can see in Figure 5 that this strategy plays a certain role in a short period of time. Although the short-term peak occurred after the date of border blockade, this is also explicable. For problems such as epidemic control, it is normal for the strategy to have a certain lag in the effectiveness, and the increased rate of confirmed cases slowed down significantly after the peak, which indicates that the strategy has played a role to some extent. However, as can be seen from the Figure 10, large-scale outbreaks of the epidemic have not been avoided in the later stage even though the border blockade has been in place throughout the whole period， which shows that although the border blockade seems to be effective in the short term, the problem of epidemic prevention and control has not been fundamentally solved due to this strategy.

For visual design, we drew two figures for the border blockade part of the third strategy, both of which showed the date of the border blockade on the graph of the number of confirmed cases changing by date. Figure 10 shows the line chart of the entire period from the beginning of the epidemic to the present and the specially marked lockdown dates, while Figure 5 shows an enlarged view of the situation about one month before and after the lockdown date.

- The horizontal and vertical coordinate layout of Figure 10 and Figure 5 is the same. The horizontal coordinate is the date, which is an ordered data, and the vertical coordinate is the number of confirmed cases in New South Wales, which is a ratio data, quantitative data.
- Only in the visual presentation, we choose different visual variables for the two graphs. For the Figure 10, the line graph is used, which has two main reasons; firstly, for the data in the whole period, the focus is to reflect the "trend" rather than the specific value, and secondly, avoid the situation that bar chart or other graphs tend to be blank without visual expression when the value is close to 0. However, we used a bar chart for Figure 5, because when the date was focused on about one month before and after the border blockade, the details were enlarged. At this time, the numerical situation of each day could be seen by using the bar chart, while the problem of low image filling rate may occur if the line chart is used instead.

PART A : VACCINATION:

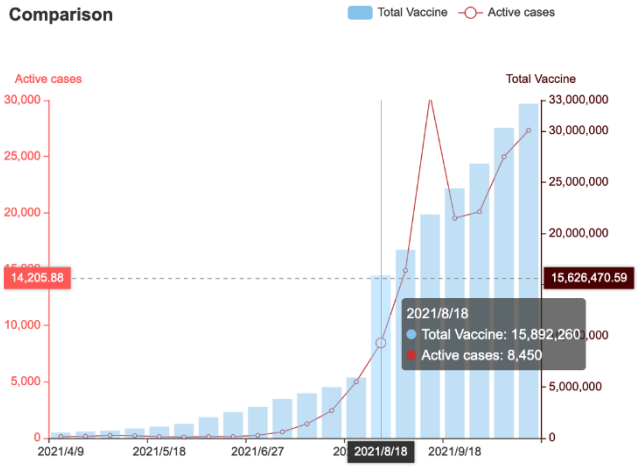

Figure 6

Vaccination is one of the effective techniques to control the spread of covid-19. The report uses a mixture of line graphs and histograms to show the impact of vaccines on the covid-19. As a result, we use line graphs to represent the changing trend of the number of confirmed daily vaccines, as well as a histogram to indicate the total number of vaccinations. The vertical axis in the figure above represents the total number of people diagnosed and the total number of people vaccinated. The abscissa of the two is the same-date. The report selected a total date of about 6 months, which is 6 months since people gradually began to realize the necessity of vaccination.

The picture uses a visual variable "color" to help distinguish the two types of data, with the line chart using a darker color and the bar chart using a lighter hue, which aggravates the color contrast and avoids unclear data in the line chart. This picture also uses an interactive design. The user can locate different data by moving the mouse.

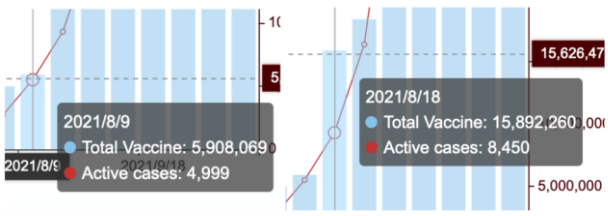

Figure 7

Figure 6 shows the change in the number of vaccinations and the total number of confirmed cases from April 19, 2021 to October 18, 2021. There is no immediate data to show that a single vaccine measure has a decisive effect on the change in the confirmed number of people due to the lag of the vaccine, which could be attributed to the following factors:

- Vaccines typically require two or three needles, and there is no way to demonstrate their effectiveness through data right away.
- Encountered external factors after the rapid increase in the number of vaccinations (the rising period of Delta virus)

However, the vaccine is still effective. We predict that the vaccine's effective time is from early September to mid-September due to the vaccine's latency. Figure 1 shows the number of active cases declined dramatically from early September to mid-September, indicating that the vaccine has a certain effect on the control of covid-19

PART B - COMBINATION OF STRATEGIES

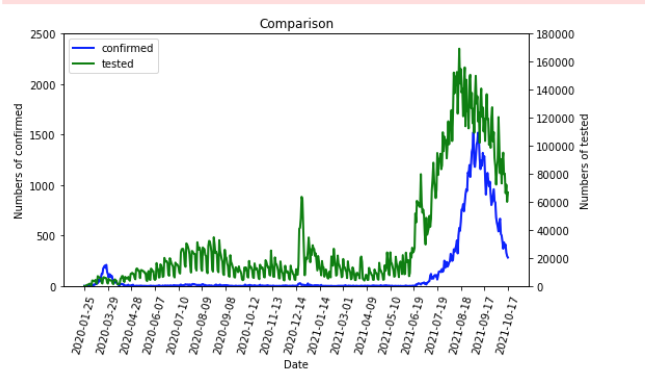

Figure 8

The figure above shows the relationship between the number of persons who have been tested and the number of people who have been diagnosed. In order to show the volatility of the number of people infected, We opted to use the x-axis as the time point when the epidemic started to October 18, and the y-axis representing the number of people diagnosed and tested. The visual variable "color" is used in this figure to help distinguish the two polylines.

As we can observe from the visualization above, there are two distinct peaks, one on March 20, 2020, and one on September 17, 2021. Obviously, the second peak was dozens of times higher than the first peak, which indicates that the measures implemented at the second peak had a significant effect.

To be more specific, the Lockdown, stay-at-home restriction and encouraging vaccines were all implemented together during that time. From the former question A Figure 8 shows the number of people vaccinated on August 18, 2021, and the Australian government states that [1] most vaccines on the market require two to three doses to provide optimal immunity, so we can estimate the effective date of the vaccine as early to mid-September 2021.

Meanwhile, during this period, the initial lockdown and the stricter stay-at-home restriction implemented on August 25 had controlled the further spread of Delta virus as well. As can be seen from the visualization above, the number of infections began to decline from mid-September which means the combination of these two measures effectively controlled the outbreak again.

PART B - EVALUATION

To find whether there's a particular strategy that works or that worked better than others, we should define our Key Goal Indicator (KGI) first. We define it as using measures to completely control the COVID19 epidemic. And what is the key performance of a particular method ? We define an important Key Performance Indicator (KPI) ourselves, that is, the number of confirmed cases in the next month after a particular measure took effect become lower than the previous month.

We give measures that can be implemented immediately (such as wearing a mask, lockdown) an effective date a lag period. We know that the number of confirmed cases cannot drop immediately the next day after the lockdown. So we define this time period as 14 days, which is the virus incubation period of COVID19. In the month after the lag period of this measure, if the number of confirmed cases becomes lower than the previous month, then we claim that the strategy has achieved the KPI.

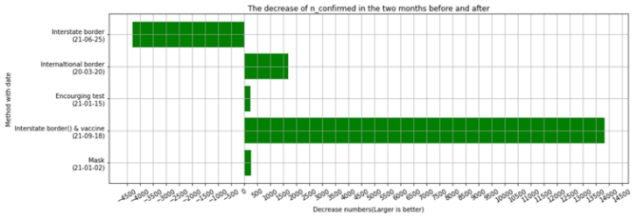

Figure 9

We create a horizontal bar chart for determining whether and how well particular methods (or combination) achieve the KPI. Decrease numbers are internal data, because it does not have true zero, the decrease numbers can be negative, which means n_confirmed increase after month. We represent these decrease numbers by the horizontal axis. They are homogeneous. As for methods(or combination), they are nominal data and heterogeneous. We arrange them on the vertical axis. Bar chart can clearly distinguish the performance of a particular method via the length of bar, and another reason why we use bar chart is that the vertical axis represents heterogeneous data.

As we can see from the chart, there's four strategies (or combination) achieve the KPI, they are Mask, combination of Interstate border lockdown and vaccine, Encouraging test, and International border lockdown. The results of the KPI are consistent with the results of our previous separate visual analysis. Moreover, it seems like a combination of strategies

consisting of interstate border and vaccine worked the best. After this combination of measures took effect, nearly 14000 confirmed cases decreased in a month. We expect this combination to completely control COVID19 in the future to achieve the KGI.

PART C: STRATEGIES DID NOT WORK AS ANTICIPATED

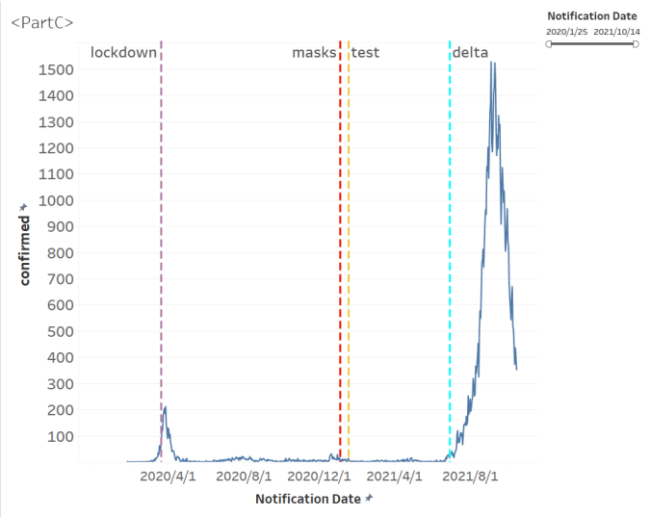

Figure 10

In order to judge which strategies did not work as expected, a line chart was selected for visual analysis. Through the line chart, the trend of the number of confirmed diagnoses over time can be seen. The horizontal axis represents time, and the vertical axis represents the number of confirmed people. The four reference lines in the figure represent different time points, and different colors can be used as visual features to facilitate the identification of different time points. In addition, a time filter is added as a visual interaction design. The specific visualization is shown in Figure 10 above.

According to the analysis before, good hygiene procedures, testing and international border can effectively control the development of the epidemic and reduce the number of confirmed cases. But as shown in Figure 10, with the emergence of delta virus on June 28, 2021, there was a clear upward trend and rapid increase in the number of confirmed cases, suggesting that these three strategies were not working as expected.

Reference:

[1] Australian Government Department of Health, "Is it true? Do people have to receive two doses and do they have to be the same type of COVID-19 vaccine for it to be effective?," *Australian Government Department of Health*, Mar. 18, 2021. https://www.health.gov.au/initiatives-and-programs/covid-19-vaccines/is-it-true/is-it-true-do-people-have-to-receive-two-doses-and-do-they-have-to-be-the-same-type-of-covid-19-vaccine-for-it-to-be-effective.

[2] data.nsw.gov.au

[3] Masahiro Takatsuka, USYD-COMP5048-WEEK8-Lecture

[4] https://matplotlib.org/stable/contents.html

[5] Masahiro Takatsuka, USYD-COMP5048-WEEK 2-Lecture

[6] Masahiro Takatsuka, USYD-COMP5048-WEEK 5,6,7-Lecture

[7] Australian Government Department of Health, "Resources | Australian Government Department of Health," *Australian Government Department of Health*, Nov. 09, 2019. https://www.health.gov.au/resources.

[8] N. G. D. Channels, "COVID-19 news and media," *NSW Government*, Jan. 02, 2021. https://www.nsw.gov.au/covid-19/news-and-media.