HARNESSING COVID-19 PUBLIC DATA FOR INFORMED HEALTH DECISIONS IN AUSTRALIA

Contents

Executive Summary

This report delivers key insights on the COVID-19 pandemic to inform the Agency for Clinical Innovation's public health strategies. Through descriptive, clustering, and correlation analyses, Various hypothesis tests, and Vaccination analysis, it maps the pandemic's trajectory and assesses intervention effectiveness both globally and within Australia.

The descriptive analysis outlines the global impact of COVID-19, noting the challenge of new variants and the benefits of vaccination campaigns. For Australia, it highlights four waves and the effects of early measures against the Delta and Omicron variants.

Clustering countries using K-Means reveals distinct pandemic impacts and response needs, essential for tailored strategies. Correlation analysis shows Australia's COVID-19 patterns are significantly aligned with some European countries, suggesting comparable trends or health responses.

Correlation analysis investigates the relationships between COVID-19 case trends in Australia and other countries. Findings show significant correlations with several European countries and regions, indicating similar pandemic patterns or responses. Furthermore, time-shifted analyses suggest potential lead times and delayed reactions in the correlation of pandemic trends between Australia and other countries.

Hypothesis testing confirms Australia's strict health measures have significantly influenced COVID-19 reproduction rates. A very weak explanatory power of the model was created; therefore, further and deeper analysis is required to reveal the effect of strict policies better by focusing on more local areas and adding variables for each measure.

Statistical tests and models, including t-tests and XGBoost, identify a marked reduction in death rates post-vaccination and pinpoint key predictors of COVID-19 mortality. OLS regression analysis further links higher vaccination rates to better health outcomes, such as reduced cases in older populations and lower ICU admissions. Public health measure assessments indicate a slight negative correlation with the COVID-19 reproduction rate, suggesting other factors are also at play.

In conclusion, the report underscores the importance of data-driven approaches in crafting public health policies. It provides ACI with a detailed assessment of the pandemic's patterns, guiding future decision-making towards a data-informed, responsive, and dynamic approach to health crises.

Introduction

This report provides an analytical overview of the COVID-19 pandemic to inform the Agency for Clinical Innovation's response strategies. It incorporates descriptive and clustering analyses, examines correlations with global trends, tests hypotheses on health measure effectiveness, and assesses the impact of vaccinations. These insights aim to enhance understanding and inform evidence-based decision-making for ongoing and future public health challenges.

Descriptive Analysis

The descriptive analysis provides a concise overview of COVID-19 trends, charting the pandemic's course through meticulous data visualisation. Focused on both Australia's response and the broader global impact, the graphs synthesize key developments, variant surges, and policy effects into a coherent visual narrative. These insights, rooted in robust datasets, are instrumental for guiding public health strategies and understanding the pandemic's multifaceted progression.

Global Overview of COVID-19 Trends
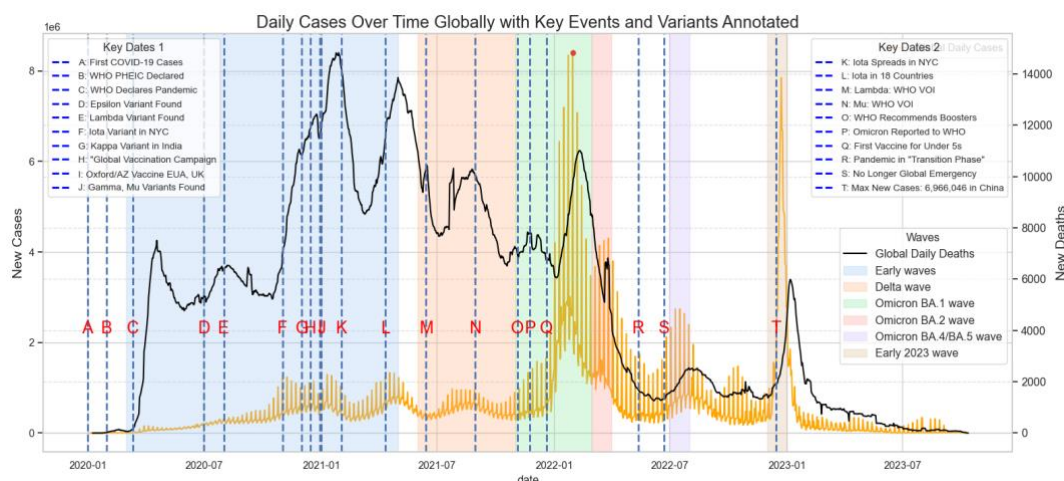
Figure 1: Panoramic view of pandemic globally with the key events and variants annotated.

The research encapsulates the global impact of COVID-19 through a meticulously crafted graph, illustrating the disease's trajectory and pivotal events. This dual-axis line

plot captures daily cases and deaths, annotated with significant developments such as variant emergence and policy milestones.

Key observations from the graph include:

- The ascent from early waves to Delta highlights the variant's increased contagion.
- The evolution from Omicron BA.1 to BA.4/BA.5 indicates ongoing viral mutation, challenging herd immunity and vaccine efficacy (World Health Organization, 2022)
- Post-vaccination, there's a noticeable decoupling of case and death rates, suggesting vaccines' role in reducing fatalities.

Policy and trend correlations in the graph suggest that global vaccination efforts like the "Global Vaccination Campaign" are aligned with drops in mortality rates, evidencing vaccine impact (Agrawal et al. 2023).

Conclusions drawn from the graph are:

- The pandemic's course has been punctuated by waves, shaped by new variants and policy responses.
- Vaccination campaigns reflect the success of global collaboration and scientific progress.
- The effectiveness of policy actions is mirrored in fluctuating case and death numbers.

The graph is pivotal for the analysis, informing policy formulation for NSW Health and emphasizing the importance of evidence-based interventions in the ongoing battle against COVID-19.

Pandemic Patterns in Australia
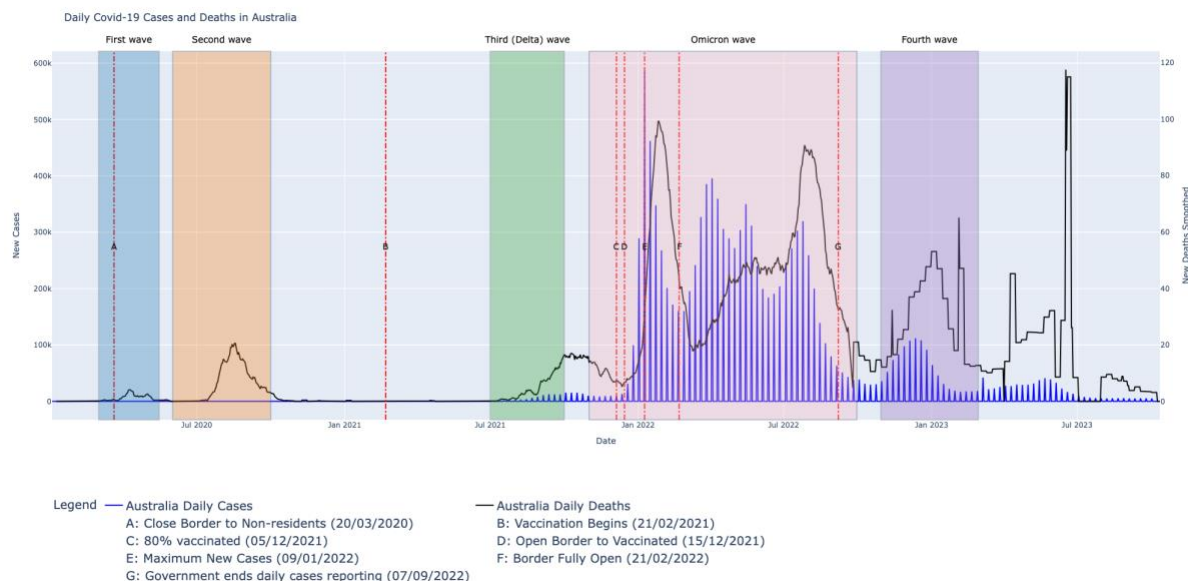
Figure 2: Panoramic View of Australia During Pandemic

In examination of the COVID-19 pandemic in Australia, we've charted four major waves: the initial outbreak, a second wave, a third wave propelled by the Delta variant, and the last dominated by Omicron, with an anticipated wave in early 2023 (Kelly, 2023). The study from the Australian Bureau of Statistics (2023), details Australia's pandemic response, including stringent border controls (Storen & Corrigan, 2020) and a robust vaccination drive.

Australia's proactive measures, such as the early 2020 border closure, resulted in low infection rates, demonstrating the effectiveness of early intervention. The Delta variant necessitated policy shifts to tackle increased transmissibility.

The 2021 vaccination campaign targeted an 80% coverage to facilitate border reopening, which was challenged by the Omicron variant's surge, escalating cases to near 60,000 daily (Hanly et al., 2021). Despite high vaccination rates, Omicron led to a significant phase with increased fatalities, a trend that persisted as daily case numbers began to fall.

This contradiction underscored the variant's lethality and its strain on healthcare systems. The government's decision to stop daily infection reporting in September 2022

(SBS News, 2022) reflects an evolving approach to pandemic management, with strategies continually adapting to new challenges.

Clustering Analysis

The purpose of this analysis was to identify distinct groups (clusters) of countries based on several features derived from Our World in Data (OWID) COVID-19 dataset. The analysis aimed to understand how countries can be grouped according to their COVID-19 cases, deaths, testing rates, vaccination rates, and other socio-economic indicators.

Data Preprocessing

The initial dataset comprised a wide range of variables. We filtered out columns related to 'new', 'smoothed', and 'weekly' statistics to focus on cumulative and static figures. Subsequently, we dropped several specific columns that were deemed less relevant for clustering or had high missing values.

We also excluded aggregate entries such as "World" and regions like continents or income groups. Missing values were a concern, especially in columns such as 'total_tests_per_thousand' and 'people_fully_vaccinated_per_hundred', which had significant gaps. These were handled by focusing on the latest available data for each country and dropping rows where data was not available.

Feature Selection and Normalisation

The remaining numeric features were standardised using the "StandardScaler" to ensure that the clustering algorithm treated all features equally, without bias towards those with larger scales.

Clustering

The K-Means algorithm was used to identify clusters. The Elbow Method was employed to find the optimal number of clusters, which was determined to be three (k=3). The Silhouette Score, which measures how similar an object is to its own cluster compared to other clusters, was calculated to be 0.319 for the chosen number of clusters, indicating a fair degree of separation.
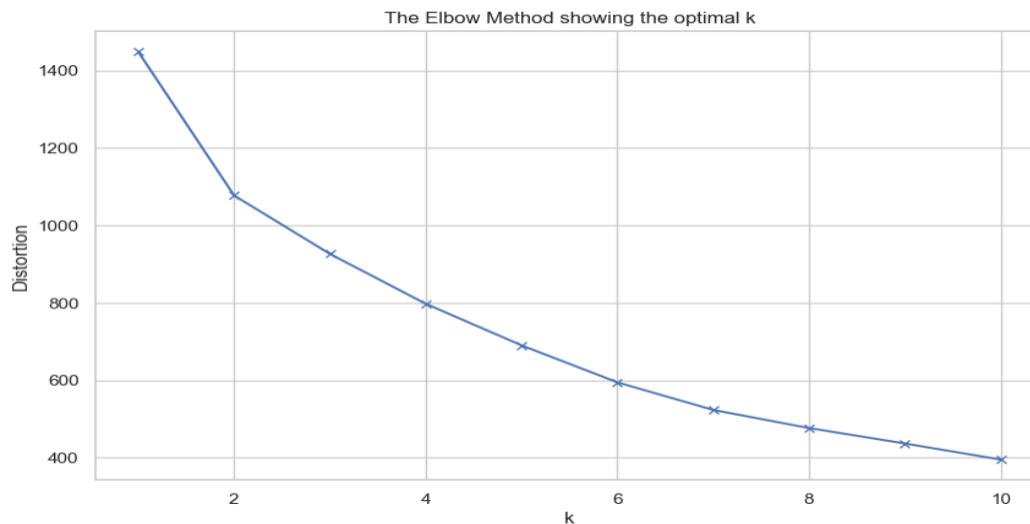
Figure 3: Elbow Method showing the optimal k.

Feature Importance

A "RandomForestClassifier" was employed to gauge the importance of each feature in the clustering process. 'GDP per capita' and 'total_cases_per_million' emerged as the most influential factors.
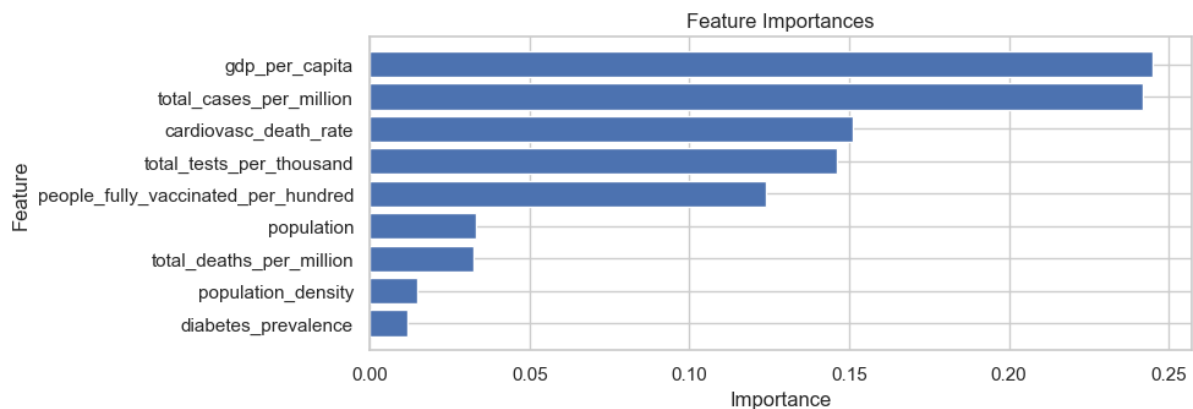
Figure 4: Feature Importance

Three distinct clusters were identified:

- Cluster 0: Predominantly high-income countries with high values for 'gdp_per_capita', 'total_cases_per_million', and 'total_tests_per_thousand'. Countries like the United States, United Kingdom, and Germany was part of this group.

- Cluster 1: A mix of middle- to low-income countries with lower 'gdp_per_capita' and 'total_tests_per_thousand'. This cluster included countries such as Afghanistan, Albania, and Algeria.
- Cluster 2: Exceptional cluster with only two countries, China and India, likely due to their unique demographic and socio-economic indicators.

Cluster Means

The means of the clusters' features further solidified the distinctions between them, with Cluster 0 having the highest mean GDP per capita, Cluster 1 the lowest, and Cluster 2 being notable for its massive population.

Visualisation of the clusters on World map

A world map visualisation indicates the geographic distribution of the clusters. It highlighted that high-income countries (Cluster 0) were primarily located in Europe, North America, and parts of Asia and Oceania. Cluster 1 countries were spread across different continents, with a significant number in Africa, Asia, and Latin America. Cluster 2 contained only China and India, each a significant population centre.
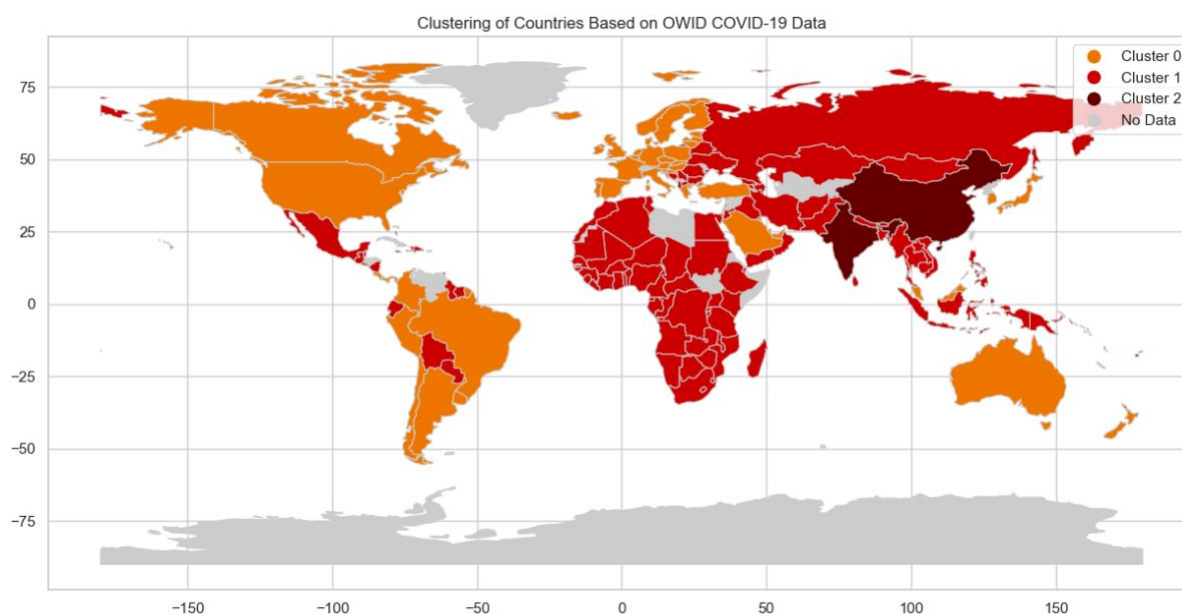
Figure 5: Cluster of Countries Based on OWID data.

The clustering analysis revealed significant variation in how countries have experienced and responded to the COVID-19 pandemic. It highlighted the influence of economic capabilities and demographic factors on the pandemic's statistics.

Policymakers and health organizations should leverage these findings to customize pandemic responses to the specific needs of different country clusters. Collaborative strategies and resource sharing among similar countries could enhance pandemic control efforts. Future research should explore changes over time, including vaccination uptake and public health policies, for a dynamic understanding of pandemic response trajectories. This analysis of the OWID COVID-19 dataset offers a framework for targeted and informed international health crisis management.

Correlation Analysis

It is critical to assess the relationship between new COVID-19 cases per million in other countries and those in Australia. Such an examination gives information on whether worldwide breakouts or containment can be used to forecast Australia's position. This might be due to a variety of causes, such as global travel patterns, economic relations, or common policy responses. Understanding this association might provide policymakers, health experts, and the public with useful information. It helps predict future case patterns in Australia based on worldwide statistics, ensuring that the country is better equipped to deal with the issues that the epidemic poses.

1. Correlation between new cases per million of other countries with Australia's

We have conducted some investigation on how many cases per million in other countries related to the same figure in Australia. According to the Figure above, we got some insight:

Portugal has the highest absolute correlation with Australia at approximately 0.715. This suggests that the trend in daily new cases per million in Portugal has been somewhat like Australia's.

Other European countries and regions (like Luxembourg, the European Union, Greece, France, Germany, Spain, Malta, and Iceland) also show significant correlation with

Australia. This could indicate similar trends or responses to the pandemic across these regions.

An unexpected entry in the list is Bonaire Sint Eustatius and Saba, which is a Caribbean region. While it's interesting to see this region on the list, it's essential to remember that correlation does not imply causation. The reasons behind similar trends could be multifactorial and may not necessarily indicate direct similarities in disease spread or response strategies.
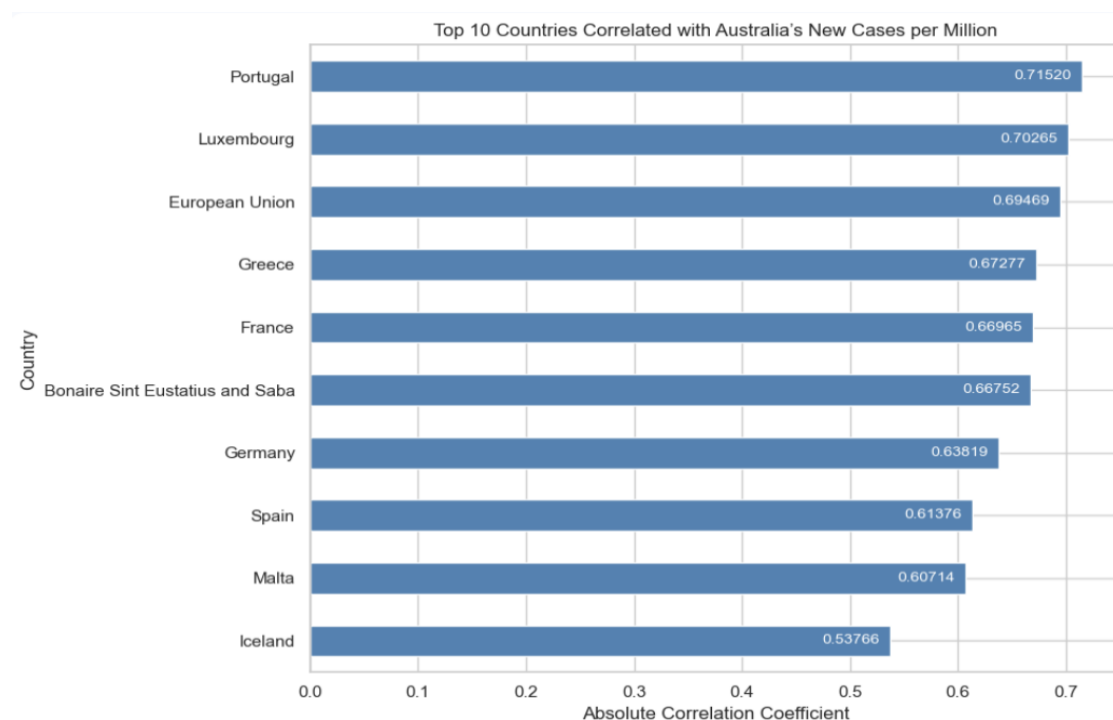
Figure 6: Top correlated countries with Australia

2. Correlation between new cases per million of other countries with Australia's with different days shift

The dynamic character of the COVID-19 pandemic, driven by many local and international causes, needs a comprehensive understanding of the links between different nations' cases over time. A straight day-to-day comparison may not convey the whole picture of how developments in one nation impact or are influenced by trends in another. In essence, although the direct correlation gives a snapshot of the

relationship, including distinct day shifts allows for a fuller view of the interplay across time. This methodology recognises the complex impacts and delayed reactions inherent in the worldwide response to the pandemic, giving greater insights into the transnational dynamics of COVID-19 and providing Australia with critical lead time for preventative action.

We decided to calculate the average correlations for each day's shift. The result is visualized in the figure below.
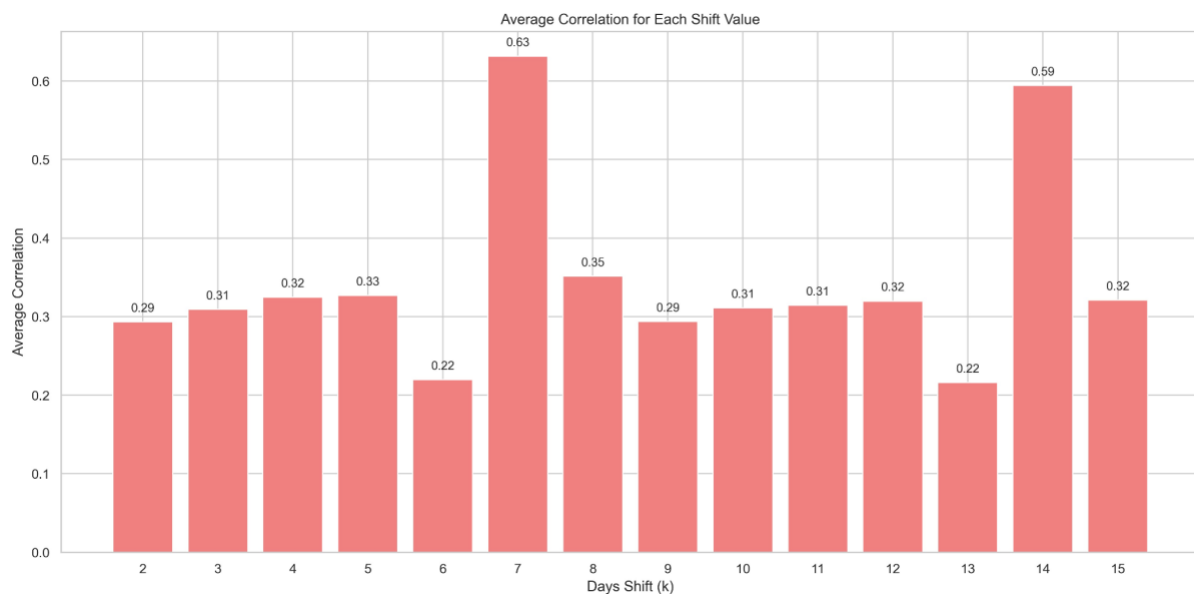
Figure 7: Average Correlation for Each Shift Value

Insights from Graph Analysis:

- Major countries do not show a consistent correlation with Australia's pandemic data, indicating unstable relationships.
- Correlated countries are globally diverse, hinting at similar pandemic trends across different regions, regardless of geographical and cultural differences.
- Shifting Australia's data by 7 days reveals high correlations with certain European countries, suggesting possible weekly reporting patterns or seasonality.
- Correlations become less consistent and vary more with shifts beyond a week.

3. Hypothesis test: Hypothesis (H1) changes in the data of each other country may be associated with changes in the data for Australia (7 days shifted)

We've established a significance threshold at 0.025. For each country, we've gauged its performance in relation to Australia. Using a standard statistical measure called the Pearson correlation coefficient, we can determine the strength and direction of the linear relationship between the two sets of data.

Alongside this, we also gauge the probability that the observed correlation occurred by chance. This is the p-value.

These insights are systematically recorded for each country, specifically noting if the relationship is above the significance benchmark. Lastly, we've provided a concise report showcasing only those countries that have a significant relationship with Australia's 7-day metrics. For these countries, the report details the strength of the correlation and the corresponding p-value.

With the result, we can conclude that:

- Many countries show significant correlations with Australia, but the magnitude and direction (positive/negative) vary.
- Some countries like Greece, France, and the European Union have very high positive correlation coefficients with Australia (all above 0.6). This implies a strong positive linear relationship between their data and Australia's data.
- A few countries, such as Hungary, Iran, Iraq, and Libya have negative correlations with Australia. Though the correlations are significant, they're not particularly strong in magnitude.
- Notably, the European Union, which is not a country but a political and economic union, also shows a strong positive correlation with Australia.

Hypothesis Testing: The Effect of Health Policy Stringency on COVID-19 Transmission in Australia

The investigation aimed to understand the relationship between the stringency of public health measures and the COVID-19 reproduction rate (R-value) in Australia. The R-

value indicates how quickly the virus is spreading: values above 1 signal an expanding epidemic, while values below 1 point to a decline.

Initial Analysis:
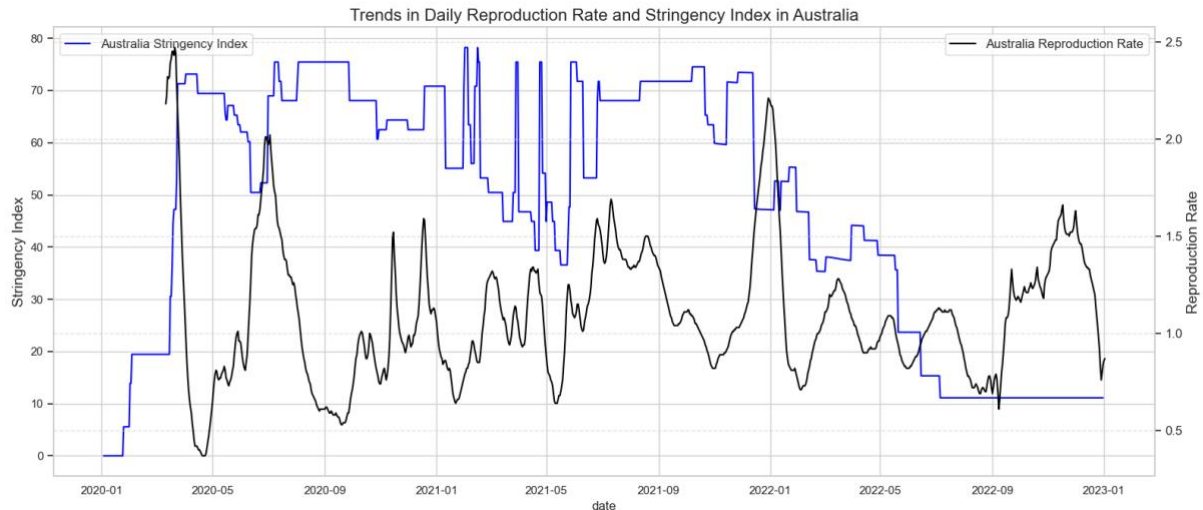
Figure 8: Trends in Daily Reproduction Rate and Stringency Index in Australia

- The plot does not reveal any correlations between the two variables.
- Pearson correlation analysis indicated a statistically significant but weak inverse relationship between stringency measures and the R-value (correlation coefficient: -0.1109; p-value: 0.00037).
- A lagged regression model suggested a weak, non-significant correlation with a low R-squared value (0.015), implying a negligible impact of prior stringency measures on the
current R-value.

```
                            OLS Regression Results
==============================================================================
Dep. Variable:       reproduction_rate   R-squared:                       0.015
Model:                             OLS   Adj. R-squared:                  0.002
Method:                  Least Squares   F-statistic:                    0.9033
Date:                 Fri, 03 Nov 2023   Prob (F-statistic):              0.549
Time:                         21:53:07   Log-Likelihood:                 -291.55
No. Observations:                 1011   AIC:                             611.1
Df Residuals:                      997   BIC:                             680.0
Df Model:                           13
Covariance Type:                   HC3
==============================================================================
                           coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const                    1.0788      0.010    105.036      0.000       1.059       1.099
lagged_diff_stringency_2d   0.0075      0.004      2.117      0.034       0.001       0.014
lagged_diff_stringency_3d   0.0062      0.003      1.795      0.073      -0.001       0.013
lagged_diff_stringency_4d   0.0046      0.003      1.381      0.167      -0.002       0.011
lagged_diff_stringency_5d   0.0045      0.003      1.410      0.158      -0.002       0.011
lagged_diff_stringency_6d   0.0034      0.003      1.073      0.283      -0.003       0.010
lagged_diff_stringency_7d   0.0030      0.003      0.919      0.358      -0.003       0.009
lagged_diff_stringency_8d   0.0026      0.003      0.780      0.435      -0.004       0.009
lagged_diff_stringency_9d   0.0016      0.003      0.484      0.629      -0.005       0.008
lagged_diff_stringency_10d  0.0010      0.003      0.290      0.772      -0.006       0.008
lagged_diff_stringency_11d  -0.0006     0.003     -0.180      0.857      -0.007       0.006
lagged_diff_stringency_12d  -0.0014     0.004     -0.398      0.690      -0.008       0.006
lagged_diff_stringency_13d  -0.0026     0.004     -0.716      0.474      -0.010       0.005
lagged_diff_stringency_14d  -0.0042     0.004     -1.065      0.287      -0.012       0.004
==============================================================================
Omnibus:                       73.817   Durbin-Watson:                   0.018
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               91.266
Skew:                           0.658   Prob(JB):                     1.52e-20
Kurtosis:                       3.661   Cond. No.                         3.38
==============================================================================

Notes:
[1] Standard Errors are heteroscedasticity robust (HC3)
```

Figure 8: OLS Regression with lags Results

Subsequent OLS Model without Lag:

An Ordinary Least Squares (OLS) regression with robust standard errors to adjust for heteroscedasticity yielded an R-squared of 0.012, showing the stringency index's limited predictive power on the R-value. This model was not satisfactory; Therefore, further actions were taken like data transformations.

Final OLS Model Analysis:

Figure 9: Final OLS Regression Results

- A refined OLS model using a log-transformed reproduction rate marginally increased the R-squared to 0.020.
- Diagnostic checks raised concerns about residual normality and independence, as indicated by the Shapiro-Wilk test and Durbin-Watson statistic, respectively.
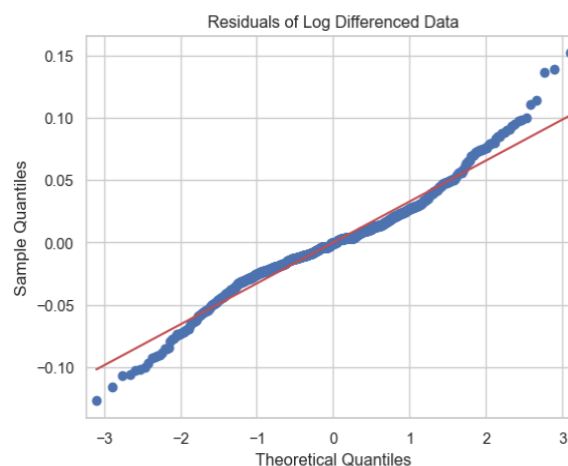
Figure 10: Residuals of Log Differenced Data

The residual analysis, including checks for linearity and equal variance, a Q-Q plot, and the Shapiro-Wilk test, revealed autocorrelation per the Durbin-Watson statistic, indicating non-independence of consecutive residuals—common in time-series data.

There is a slight negative correlation between public health measure stringency and COVID-19's R-value, but the model's low explanatory power suggests significant other factors at play.

In summary, there is a negative association between policy stringency and the COVID-19 reproduction rate, but the low R-squared value cautions that additional variables likely influence the R-value. Future analysis should consider confounding factors, utilize more detailed models, and examine local variations for clearer insights.

Vaccination Analysis

Method

Extreme Gradient Boosting (XGB), Shapley Additive exPlanations (SHAP) and Ordinary Least Squares regression (OLS)

Vaccination distribution
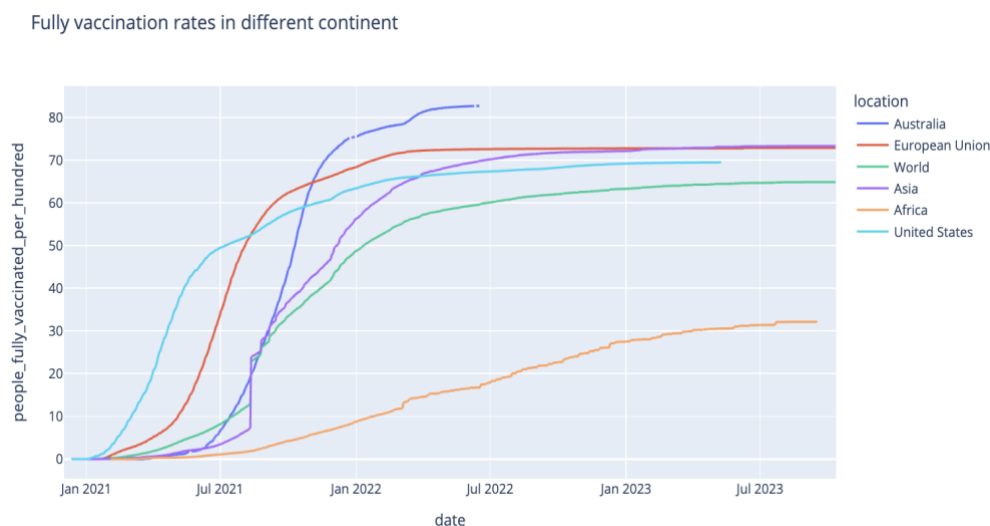
Figure 11: Fully vaccinated rates in different continents

Illustrated in Figure 8, fully vaccination coverage of 80% was achieved in Australia within a mere 6 months, displaying the most rapid ascent, characterised by the steepest of the line compared with other regions.
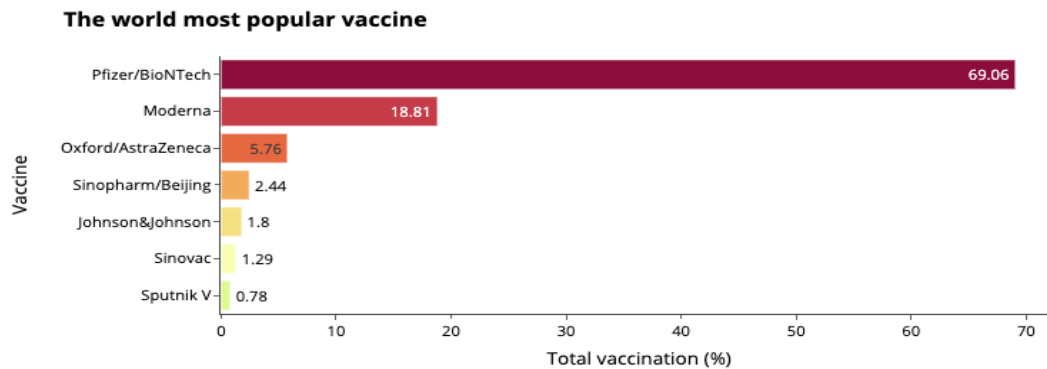
Figure 12: Popular vaccine used in the world.

As we can see in this figure, it provides a snapshot of the most commonly used vaccines worldwide. Based on the available data of WHO, Pfizer and Moderna vaccines are widely administered globally. These vaccines have played a crucial role against COVID-19.

1. **Pre-vaccination and post-vaccination effects on death rate**

- H0: There is no significant difference in death rates before and after the vaccination campaign implementation.

- H1: There is a significant difference in death rates before and after the vaccination campaign implementation.

The paired sample t-test is performed to compare death rates before and after vaccination. The result indicates that significant difference in death rates before and after vaccination (p-value < 0.05). The data supports the effectiveness of vaccination in reducing mortality rates. It signifies pivotal evidence towards achieving a more secure community, and increase confidence in vaccination campaigns.

2. Vaccination impact on death rate
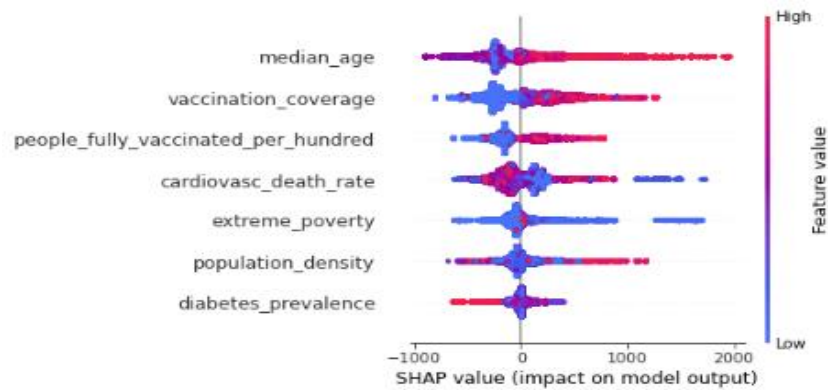
Figure 13: Feature importance on predicting death rate.

The XGB model is trained using features related to vaccination coverage, full vaccination rate, extreme poverty rate, population density, and the two underlying health indicators including cardiovascular effect and diabetes prevalence to predict death rates. Root Mean Squared Error (RMSE) equals approximately 258.86 which is considered the XGB model is appropriate. As in Figure 10, features with higher absolute SHAP values have a greater impact on the predictions. It helps in understanding the direction and magnitude of the effect. A low full vaccination rate increases the death rate when the COVID-19 wave comes up.
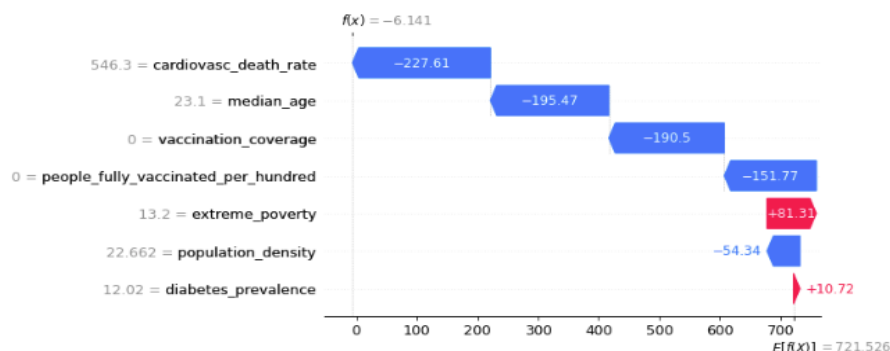
Figure 14: Waterfall plot of weighted coefficients of features effects on death rate

Feature contributions

In Figure 11, the feature cardiovascular death rate has the largest contribution around 227.61. A higher rate of cardiovascular underlying health condition cases is associated with the predicted total deaths per million. The age feature has the second-largest contribution around 195.47. A higher median age is associated with a decrease in the predicted total deaths per million. Vaccination coverage has a third-largest negative contribution approximately 190.5. A higher vaccination coverage is associated with a decrease in the predicted total deaths per million. Fully vaccinated rate has a negative contribution 151.77 which indicates that higher fully vaccination and booster usage is associated with a decrease in the predicted total deaths per million. The extreme poverty indicators have a positive contribution 81.31 demonstrates higher value of poverty is associated with an increase in the death rate. The diabetes prevelance has a positive contribution 10.72 shows that higher value of diabetes prevalence is associated with an increase in the predicted total deaths per million. These nuanced contributions shed light on factors that significantly impact death rates. Addressing underlying health conditions and socioeconomic factors is essential in designing targeted vaccination strategies.

3. Vaccination impacts on age group, underlying health conditions and health outcomes

Several OLS regression analyses were conducted to understand the relationship between vaccination rates and various health indicators.

Impact of vaccination on age group greater than 70

```
                          OLS Regression Results
==============================================================================
Dep. Variable:          aged_70_older   R-squared:                      0.015
Model:                            OLS   Adj. R-squared:                 0.015
Method:                 Least Squares   F-statistic:                    5458.
Date:                Tue, 24 Oct 2023   Prob (F-statistic):              0.00
Time:                        13:56:47   Log-Likelihood:            -1.0006e+06
No. Observations:              348538   AIC:                        2.001e+06
Df Residuals:                  348536   BIC:                        2.001e+06
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                                  coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const                           3.8222      0.010    394.780      0.000       3.803       3.841
people_fully_vaccinated_per_hundred  0.0168  0.000     73.875      0.000       0.016       0.017
==============================================================================
Omnibus:                    41392.749   Durbin-Watson:                  0.001
Prob(Omnibus):                  0.000   Jarque-Bera (JB):           58315.806
Skew:                           1.001   Prob(JB):                        0.00
Kurtosis:                       3.093   Cond. No.                        57.0
==============================================================================
```

Figure 9: OLS regression for relationship vaccination on cases of age group greater than 70

- H0: There is no significant impact of vaccination on elderly cases.
- H1: Vaccination has a significant impact on elderly cases.

The result of the OLS model suggests rejecting H0 and accept H1. The model demonstrates a significant relationship between vaccination and cases among the elderly (age group greater than 70). This underscores the importance of prioritising vaccinations for this vulnerable demographic.

Impact of vaccination on life expectancy

```
                          OLS Regression Results
==============================================================================
Dep. Variable:         life_expectancy   R-squared:                    0.000
Model:                             OLS   Adj. R-squared:               0.000
Method:                  Least Squares   F-statistic:                  93.12
Date:                Tue, 24 Oct 2023   Prob (F-statistic):         4.96e-22
Time:                        13:56:48   Log-Likelihood:           -1.5593e+06
No. Observations:              348538   AIC:                       3.119e+06
Df Residuals:                  348536   BIC:                       3.119e+06
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                                  coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const                          67.5080      0.048   1403.765      0.000      67.414      67.602
people_fully_vaccinated_per_hundred  0.0109    0.001      9.650      0.000       0.009       0.013
==============================================================================
Omnibus:                   173701.045   Durbin-Watson:                0.001
Prob(Omnibus):                  0.000   Jarque-Bera (JB):        807764.357
Skew:                          -2.539   Prob(JB):                      0.00
Kurtosis:                       8.462   Cond. No.                      57.0
==============================================================================
```

Figure 10: OLS regression for relationship vaccination on life expectancy

- H0: There is no significant impact of vaccination on life expectancy.
- H1: Vaccination has a significant impact on life expectancy.

The result of the OLS model suggests rejecting H0 and accept H1. This analysis aims to determine that there is a statistically significant relationship between the vaccination rate and life expectancy. This indicates that widespread vaccination efforts contribute positively to overall public health.

Impact of vaccination on cardiovascular death rate

```
                    OLS Regression Results
==============================================================================
Dep. Variable:     cardiovasc_death_rate   R-squared:                    0.038
Model:                               OLS   Adj. R-squared:               0.038
Method:                    Least Squares   F-statistic:              1.379e+04
Date:                   Tue, 24 Oct 2023   Prob (F-statistic):            0.00
Time:                           13:56:48   Log-Likelihood:          -2.2418e+06
No. Observations:                 348538   AIC:                      4.484e+06
Df Residuals:                     348536   BIC:                      4.484e+06
Df Model:                              1
Covariance Type:               nonrobust
==============================================================================
                                   coef   std err       t     P>|t|    [0.025    0.975]
------------------------------------------------------------------------------
const                            231.5400   0.341   679.380   0.000   230.872   232.208
people_fully_vaccinated_per_hundred  -0.9394   0.008  -117.425   0.000    -0.955    -0.924
==============================================================================
Omnibus:                        6502.482   Durbin-Watson:                0.002
Prob(Omnibus):                     0.000   Jarque-Bera (JB):          6708.242
Skew:                              0.327   Prob(JB):                      0.00
Kurtosis:                          2.818   Cond. No.                      57.0
==============================================================================
```

Figure 11: OLS regression for relationship vaccination on cardiovascular death rate

- H0: There is no significant impact of vaccination on cardiovascular death rate.
- H1: Vaccination has a significant impact on cardiovascular death rate.

The result of the OLS model suggests rejecting H0 and accept H1. The analysis reveals a statistically significant link between vaccination rates and reduced cardiovascular death rates. This highlights the potential correlation between vaccination rate and underlying health conditions.

Impact of vaccination on ICU patients

```
                    OLS Regression Results
==============================================================================
Dep. Variable:      icu_patients_per_million   R-squared:                  0.000
Model:                                  OLS   Adj. R-squared:             -0.000
Method:                       Least Squares   F-statistic:               0.02206
Date:                      Tue, 24 Oct 2023   Prob (F-statistic):          0.882
Time:                              13:56:48   Log-Likelihood:          -1.2937e+06
No. Observations:                    348538   AIC:                      2.587e+06
Df Residuals:                        348536   BIC:                      2.587e+06
Df Model:                                 1
Covariance Type:                  nonrobust
==============================================================================
                                      coef    std err       t     P>|t|    [0.025    0.975]
------------------------------------------------------------------------------
const                               2.1396     0.022   95.316   0.000    2.096     2.184
people_fully_vaccinated_per_hundred -7.826e-05  0.001   -0.149   0.882   -0.001     0.001
==============================================================================
Omnibus:                       461404.406   Durbin-Watson:                0.002
Prob(Omnibus):                      0.000   Jarque-Bera (JB):       87447541.132
Skew:                               7.659   Prob(JB):                      0.00
Kurtosis:                          79.072   Cond. No.                      57.0
==============================================================================
```

Figure 12:OLS regression for relationship vaccination on ICU patients

- H0: There is no significant impact of vaccination on ICU patients per million.
- H1: Vaccination has a significant impact on ICU patients per million.

The result of the OLS model suggests rejecting H0 and accept H1. The model strongly supports the notion that vaccination plays a crucial role in reducing the number of ICU patients per million. This suggests a substantial decrease in severe cases.

Future considerations

These findings emphasise the paramount importance of vaccination in shaping public health outcomes. For businesses, this underscores the importance of supporting vaccination efforts. A vaccinated workforce could lead to fewer absences, improved productivity, and more secure work environment which is known as herd immunity. Additionally, businesses in the healthcare sector might experience reduced strain on resources and facilities. As vaccinations continue, ongoing monitoring and adaptation of strategies are crucial. Businesses should remain agile, aligning policies with evolving public health interventions.

Conclusion

The report highlights the critical role of data-driven analysis in managing the COVID-19 pandemic. Key findings include the effectiveness of Australia's vaccination campaign and the importance of bespoke public health strategies informed by clustering and correlation analyses with global counterparts. The analysis indicates that while Australia's strict measures have impacted virus transmission, the complexity of the pandemic's dynamics necessitates a more localised and nuanced analysis for policy formulation.

Statistical methods have shown a clear benefit of vaccinations in reducing mortality rates and an association between higher vaccination rates and improved health outcomes. However, a multifaceted approach is needed, as public health measures alone show a limited correlation with changes in the reproduction rate of the virus.

For the Agency for Clinical Innovation, the path forward involves continued reliance on robust data analysis to refine and adapt public health strategies. As the pandemic evolves with the agility to apply lessons learned to future health emergencies. By maintaining a dynamic and evidence-based policy framework, we can ensure that public health interventions are as effective and responsive as possible.

References

Agrawal, V., Sood, N., & Whaley, C. M. (2023). The impact of the global Covid-19 vaccination campaign on all mortality. *NBER Working Paper No. 31812.* https://www.nber.org/papers/w31812

Australian Bureau of Statistics. (2022). *COVID-19 Mortality by wave.* https://www.abs.gov.au/articles/covid-19-mortality-wave

Australia's COVID-19 cases and death numbers will no longer be published daily. (2022, September 7). *SBS News.* https://www.sbs.com.au/news/article/australias-covid-19-cases-and-death-numbers-will-no-longer-be-published-daily-heres-why/pw8reh1le

Hanly, M, J., Churches, T., Fitzgerald, O., Post, J, J., MacIntyre, C, R., & Jorm, L. (2021). The impact of re-opening the international border on COVID-19 hospitalization in Australia: a modelling study. *The Medical Journal of Australia, 216(*1*),* 39-42. https://doi.org/10.5694/mja2.51291

Kelly, P. (2023). Lessons from the fourth Omicron COVID-19 wave. *Australian Government Department of Health and Aged Care.* https://www.health.gov.au/resources/publications/lessons-from-the-fourth-omicron-covid-19-wave-chief-medical-officer-professor-paul-kelly

One year since the emergence of COVID-19 virus variant Omicron. (2022, November 25). *World Health Organization.* https://www.who.int/news-room/feature-stories/detail/one-year-since-the-emergence-of-omicron

Storen, R., & Corrigan, N. (2020). COVID-19: a chronology of state and territory government announcements. *Parliament of Australia.* https://www.aph.gov.au/About_Parliament/Parliamentary_departments/Parliamentary_Library/pubs/rp/rp2021/Chronologies/COVID-19StateTerritoryGovernmentAnnouncements