

# COVID-19 Vaccination Impact Analysis

**NAME: KRISHNA VAMSI UPPALA**

## **Dataset:**

The dataset selected for the research paper is 'COVID-19' : 'owid-covid-data.csv'. Link to download the dataset: <https://covid.ourworldindata.org/data/owid-covid-data.csv>

**Research Question:** How do changes in COVID-19 vaccination rates correlate with the changes in new cases and deaths among different countries for the year 2022?

## **Introduction and Research Problem:**

The outbreak of COVID-19 in late 2019 has had systemic repercussions affecting health and socio-economic systems worldwide. Vaccination has been crucial in combating the pandemic mainly by decreasing the number of new infections and fatalities. However, as the COVID-19 vaccine has been available worldwide, the patterns of the new cases and deaths in relation to the vaccination progress differ from one country to another. This research seeks to fill this gap by assessing the variations in COVID-19 vaccination rates and how it affects the number of new cases and deaths recorded in different countries in the year 2022.

The primary research question addresses the vaccination rates of COVID-19 and their effect on new cases and deaths. It is imperative to recognize this relationship in order to plan a proper vaccination scheme and to create proper public health policies.

## **Research Questions:**

1. How do changes in COVID-19 vaccination rates correlate with changes in new cases among different countries in 2022?
2. How do changes in COVID-19 vaccination rates correlate with changes in deaths among different countries in 2022?

## **Research Objectives:**

1. To analyse the correlation between vaccination rates and new COVID-19 cases in various countries.
2. To examine the relationship between vaccination rates and COVID-19-related deaths globally.
3. To identify patterns and anomalies in the data that could provide insights into the effectiveness of vaccination campaigns.

### **Potential Impact of the research:**

The importance of this study is in its contribution to the knowledge that can be used in designing public health policies and vaccinations. Analysing the relationship between the vaccination rates and COVID-19 cases and deaths, the governments and health organisations can use the results to better time their vaccination processes to minimise the impact of the virus. The results of these studies could also benefit the academic discipline by providing insights missing in the current literature on the efficacy of the COVID-19 vaccines across different socio-economic and geographical environments.

### **Literature Review:**

#### **Relevant Theories and Previous Research**

Herd immunity and the immunity obtained after vaccinating are the theoretical frameworks on which this study is based. A review of Ioannidis (2020) and So and Woo (2020) shows that governments need to work collectively to ensure that vaccines are made available to everyone, and that achieving herd immunity can be beneficial in the long run. It is important to note that our results can also be placed into the context of theoretical models of disease/vaccination as discussed for instance by Plotkin et al. (2017).

#### **Analysis of Findings and Methodologies of Past Studies:**

Past published research has shown the effectiveness of COVID-19 vaccines in trial and post-vaccination trials. For example, Baden et al. (2021) examined the feasibility of the Moderna vaccine while Polack et al. (2020) assessed the feasibility of the Pfizer-BioNTech vaccine. RCT was the common method employed in these studies to determine the effectiveness of the vaccines in preventing symptomatic COVID-19. In a similar vein, the effectiveness of the Pfizer-BioNTech vaccine was evaluated in a real-world study by Dagan et al. (2021) and Haas et al. (2021) where it was found that COVID-19 cases and death rate reduced dramatically after the vaccination campaign started in Israel.

#### **Gaps in the Existing Literature:**

While these studies shed much-needed light, there is a gap in literature on the relationship between vaccination rates and changes in cases and deaths due to COVID-19 across countries in 2022. Much of the literature focuses on primarily establishing the efficacy of the vaccine in its early days and its short-term effects in local regions. There is an urgent need to adopt comprehensive analyses at the global level that respect the socio-economic and health-related parameters that differ and thus influence the outcome of vaccination.

## Methodology

This research will employ regression-based analysis to examine the correlation between vaccination rates and changes in COVID-19 cases and deaths. Specifically, we will use multiple linear regression and time-series analysis to account for various factors influencing these outcomes.

### Target Variables:

The selected target variables are 'changes\_in\_new\_cases\_smoothed' and 'changes\_in\_new\_deaths\_smoothed' - (produced from new\_cases\_smoothed.diff() and new\_deaths\_smoothed.diff()) (Note : values are calculated per million)

### Input Variables:

The selected input variables are : 'iso\_code', 'continent', 'location', 'date', 'reproduction\_rate', 'icu\_patients', 'icu\_patients\_per\_million', 'hosp\_patients', 'hosp\_patients\_per\_million', 'weekly\_icu\_admissions', 'weekly\_icu\_admissions\_per\_million', 'weekly\_hosp\_admissions', 'weekly\_hosp\_admissions\_per\_million', 'total\_tests', 'new\_tests', 'total\_tests\_per\_thousand', 'new\_tests\_per\_thousand', 'new\_tests\_smoothed', 'new\_tests\_smoothed\_per\_thousand', 'positive\_rate', 'tests\_per\_case', 'tests\_units', 'total\_vaccinations', 'people\_vaccinated', 'people\_fully\_vaccinated', 'total\_boosters', 'new\_vaccinations', 'new\_vaccinations\_smoothed', 'total\_vaccinations\_per\_hundred', 'people\_vaccinated\_per\_hundred', 'people\_fully\_vaccinated\_per\_hundred', 'total\_boosters\_per\_hundred', 'new\_vaccinations\_smoothed\_per\_million', 'new\_people\_vaccinated\_smoothed', 'new\_people\_vaccinated\_smoothed\_per\_hundred', 'stringency\_index', 'population\_density', 'median\_age', 'aged\_65\_older', 'aged\_70\_older', 'gdp\_per\_capita', 'extreme\_poverty', 'cardiovasc\_death\_rate', 'diabetes\_prevalence', 'female\_smokers', 'male\_smokers', 'handwashing\_facilities', 'hospital\_beds\_per\_thousand', 'life\_expectancy', 'human\_development\_index', 'population', 'excess\_mortality\_cumulative\_absolute', 'excess\_mortality\_cumulative', 'excess\_mortality', 'excess\_mortality\_cumulative\_per\_million'.

This research uses regression-based analysis in line with its objectives since it can easily assess several variables and their interdependencies. It can find out considerable predictor variables and assess the magnitude of their effects on the predefined target variables. In the analysis part, the time series approach will be used as it also enables the identification of temporal patterns in the data and, therefore, gives a detailed picture of the relationship between the level of vaccination and the COVID-19 impact.

## Steps for Data Analysis:

### Data Preprocessing:

- Cleaning and formatting the dataset
- Handling missing values
- Standardising variables as needed **Exploratory Data Analysis (EDA):**
- Generating descriptive statistics
- Visualising data distributions and correlations
- Identifying outliers and anomalies

### Regression Analysis:

- Building multiple linear regression models
- Evaluating model performance using metrics such as R-squared and RMSE
- Conducting residual analysis to check for model assumptions

### Time-Series Analysis:

- Decomposing time-series data into trend, seasonality, and residual components.
- Applying autoregressive integrated moving average (ARIMA) models to analyze temporal patterns

## Data Analysis Plan - Exploratory Data Analysis (EDA)

### Descriptive Statistics of the target variables:

	changes_in_new_cases_smoothed	changes_in_new_deaths_smoothed
count	92684	92684
mean	0.3e-4	2.37e-7
std	337.12	1.0409
min	-32696.004	-129.488
25 %	0	0
50 %	0	0
75 %	0	0

max	33565.213	129.488
-----	-----------	---------

The standard deviations indicate that changes in new cases have a much wider spread compared to changes in new deaths and the range is quite large for both variables, showing extreme values in both directions (both large increases and decreases).

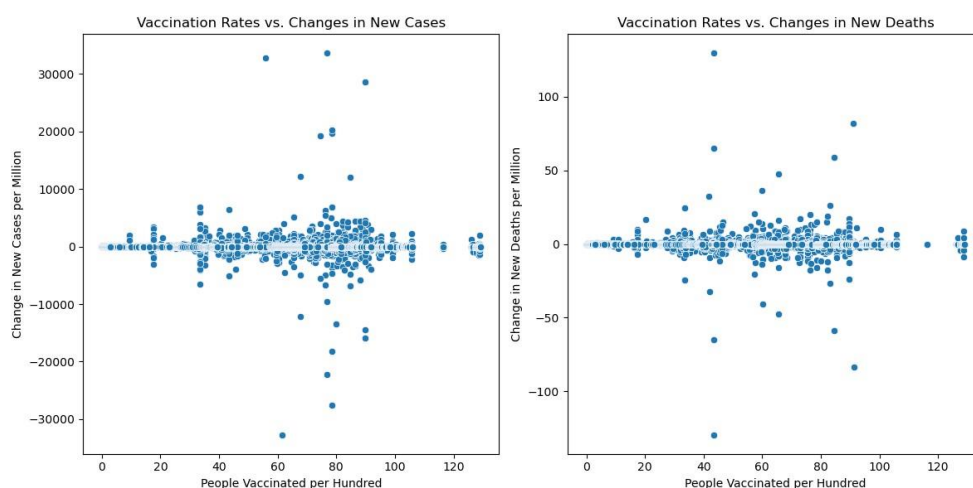
**Histograms:** To illustrate the vaccination rates and COVID-19 results in order to represent them.

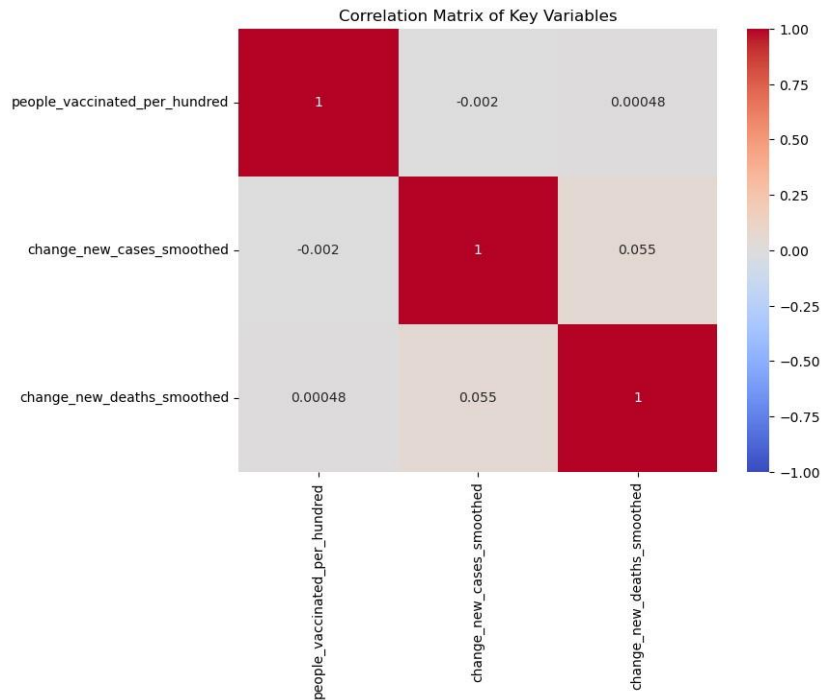
**Scatter Plots:** As an example of correlation analysis to linking vaccination density and swings in new cases, as well as deaths.

**Time-Series Plots:** To plot PESTEL analysis for better understanding of current trends and patterns of the market.

Through the implementation of the said techniques and graphics, a holistic assessment of the relationship between COVID-19 vaccinations and fluctuations in new cases and deaths shall be achieved, which will benefit public health approaches and scholarly works.

### Example Visualisations :





## References:

1. Ioannidis, J. P. A. (2020). Coronavirus disease 2019: The harms of exaggerated information and non-evidence-based measures. *European Journal of Clinical Investigation*, 50(4), e13222.
2. So, A. D., & Woo, J. (2020). Reserving coronavirus disease 2019 vaccines for global access: Cross sectional analysis. *BMJ*, 371, m4750.
3. Plotkin, S. A., Orenstein, W. A., & Offit, P. A. (2017). *Vaccines* (7th ed.). Elsevier.
4. Baden, L. R., El Sahly, H. M., Essink, B., Kotloff, K., Frey, S., Novak, R., ... & Zaks, T. (2021). Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. *New England Journal of Medicine*, 384(5), 403-416
5. Polack, F. P., Thomas, S. J., Kitchin, N., Absalon, J., Gurtman, A., Lockhart, S., ... & Gruber, W. C. (2020). Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. *New England Journal of Medicine*, 383(27), 2603-2615.
6. Dagan, N., Barda, N., Kepten, E., Miron, O., Perchik, S., Katz, M. A., ... & Hernán, M. A. (2021). BNT162b2 mRNA Covid-19 vaccine in a nationwide mass vaccination setting. *New England Journal of Medicine*, 384(15), 1412-1423
7. Haas, E. J., Angulo, F. J., McLaughlin, J. M., Anis, E., Singer, S. R., Khan, F., ... & Alroy-Preis, S. (2021). Impact and effectiveness of mRNA BNT162b2 vaccine against SARS-CoV-2 infections and COVID-19 cases, hospitalizations, and deaths following a nationwide vaccination campaign in Israel. *Lancet*, 397(10287), 1819-1829