World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:17, No:7, 2023

Analyzing Factors Impacting COVID-19 Vaccination Rates

Dongseok Cho, Mitchell Driedger, Sera Han, Noman Khan, Mohammed Elmorsy, Mohamad El-Hajj

Abstract—Since the approval of the COVID-19 vaccine in late 2020, vaccination rates have varied around the globe. Access to a vaccine supply, mandated vaccination policy, and vaccine hesitancy contribute to these rates. This study used COVID-19 vaccination data from Our World in Data and the Multilateral Leaders Task Force on COVID-19 to create two COVID-19 vaccination indices. The first index is the Vaccine Utilization Index (VUI), which measures how effectively each country has utilized its vaccine supply to doubly vaccinate its population. The second index is the Vaccination Acceleration Index (VAI), which evaluates how efficiently each country vaccinated their populations within their first 150 days. Pearson correlations were created between these indices and country indicators obtained from the World Bank. Results of these correlations identify countries with stronger Health indicators such as lower mortality rates, lower age-dependency ratios, and higher rates of immunization to other diseases display higher VUI and VAI scores than countries with lesser values. VAI scores are also positively correlated to Governance and Economic indicators, such as regulatory quality, control of corruption, and GDP per capita. As represented by the VUI, proper utilization of the COVID-19 vaccine supply by country is observed in countries that display excellence in health practices. A country's motivation to accelerate its vaccination rates within the first 150 days of vaccinating, as represented by the VAI, was largely a product of the governing body's effectiveness and economic status, as well as overall excellence in health practises.

Keywords—Data mining, Pearson Correlation, COVID-19, vaccination rates, hesitancy.

I. Introduction

CORONAVIRUS disease (COVID-19) is a highly infectious disease known to affect the respiratory and cardiovascular systems. According to the World Health Organization (WHO), most people infected with the virus experience mild respiratory symptoms and do not require medical attention. However, individuals with underlying medical conditions and those older have an increased risk of serious illness. Staying informed about the disease, practising social distancing, wearing a proper mask and staying sanitized are all methods to slow the transmission of the disease [1]. COVID-19 began in Wuhan, China, in December of 2019, where a pneumonia outbreak raised international concern [2]. Before long, the virus spread to most countries and was starting to have a massive global impact. As 2020 closed, the world dramatically changed, as COVID-19 greatly impacted people's livelihood, public health, and the workforce. Countries enforced varying responses to the pandemic, including physical interaction lockdowns, travel restrictions, and masking laws. Almost half of the world's population was at risk of losing their livelihood [3].

D. Cho, M. Drieger, S. Han, N. Khan, M. Elmorsy and M. El-Hajj are with Department of Computer Science, MacEwan University, Canada (e-mail: elhajjm@macewan.ca).

As the COVID-19 virus spread throughout 2020, different pharmaceutical companies have been developing vaccines. By December, the first vaccinations were approved for public use, and individuals were starting to get vaccinated [4]. Companies that developed the vaccine during this time include Pfizer, Moderna, AstraZeneca, and Johnson & Johnson.

The COVID-19 vaccine was developed at a rapid pace, as vaccines often take many years to be developed [4]. Multiple factors that influenced the distribution and administering of the vaccine were variable in each country. When a country first began to vaccinate was likely a product of how quickly they could acquire vaccination doses in the masses. For example, lower-income countries in Africa do not typically have the means to manufacture vaccinations locally [4]. Dependency on imported vaccine doses slows the advancement of vaccination rates. Another factor that impacted the vaccination rate was the population's hesitancy to vaccinate. The SAGE working group under the World Health Organization defines vaccine hesitancy as the refusal and/or delay of vaccination despite access to vaccination services [5]. A study done in late 2020 found a large variance of COVID-19 vaccine hesitancy among different countries [6].

Due to factors such as these, analyzing COVID-19 vaccination rates alone does not entirely represent their political and economic context. While the countries that first started vaccinating were able to make earlier headway in COVID-19 case reduction, they had to develop distribution strategies which late-starting countries could apply to their own vaccination efforts. Different start dates, populations, wealth, and access have substantially impacted the vaccine rollout. Two indices were created from the data collected in this study to properly represent unique viewpoints of each country's vaccination efforts: the Vaccine Utilization Index (VUI) and the Vaccination Acceleration Index (VAI). Both indices rank each country on a scale from 1-100 and are impacted by variables such as the date of first vaccination. The VUI accounts for supply usage and vaccinated percentages of the population to score how effectively a country utilized its vaccine supply throughout the pandemic. The VAI ranks countries on their vaccination rates 150 days following their first vaccination date, measuring how effectively countries rolled out their vaccines once they obtained the supply. This study analyzes the relationships between these two indices and country indicators.

II. Related Literature

The academic environment surrounding COVID-19 is constantly developing as the global effects of the virus are observed.

A novel coronavirus outbreak of global health concern is one of the earliest publications regarding the COVID-19 pandemic. This literature describes the early onset of the virus, beginning in Wuhan, China, in late 2019. It details the clinical features of the first patients that were confirmed to be infected with COVID-19, and out of the 41 patients in the first cohort of testing, 22 developed dyspnoea, 13 had to be admitted to an intensive care unit, and six died. While these findings were stated to be treated cautiously, such a high mortality rate was alarming. The study serves as a warning, with the ending statement calling for all efforts to understand the disease and act against it [2]. This study serves as foundational knowledge of the COVID-19 pandemic.

How COVID-19 vaccine supply chains emerged during a pandemic is a comprehensive look at the formation of COVID-19 vaccine supply chains concerning different countries and companies. It recalls how the vaccine supply chains under Pfizer, Moderna, Novavax and other companies were developed between 2020 and 2021. The vaccine distribution rolled out differently in different world areas; for example, the United States and the European Union had higher rates of mRNA vaccines such as Pfizer and Moderna being administered, while China strictly administered domestic vaccines. The study concludes by encouraging researchers to determine if there was a quicker way to manufacture these vaccines in order to prepare for the next pandemic [4]. While this study extensively highlights how the vaccine supply was distributed, it would benefit from the vaccine utilization conclusions that are found later in this paper.

COVID-19 Vaccine Hesitancy Worldwide: A Concise Systematic Review of Vaccine Acceptance Rates summarizes COVID-19 vaccine hesitancy and displayed the findings of 31 published studies. A large variety of acceptance rates were found in this study. Countries with the highest vaccine acceptance rates include Ecuador, Malaysia and Indonesia, with over 90% acceptance, while Kuwait, Jordan and Italy showed high hesitancy rates. Some surveys were among healthcare workers only, with the acceptance range from 78.1% to 27.7%. Lower acceptance rates were observed throughout the Middle East, Russia, Africa, and some European countries. This study was completed in January 2021, and it concludes by encouraging governments and other bodies to build vaccine trust with the general public in order to prevent vaccine hesitancy from hindering rates of COVID-19 vaccination [6]. Combining this paper's findings with the under-utilization of COVID-19 vaccines that some countries displayed may yield meaningful results.

A Global Study on the Correlated of Gross Domestic Product (GDP) and COVID-19 Vaccine Distribution demonstrates the association between the GDP of various countries and the progress of COVID-19 vaccinations. This study determined that wealthier nations have higher vaccination rates. It also determined that the COVID-19 vaccine rollout is uneven; countries in the Global South are not performing as well as those in the Global North, proving significant country-level disparities in achieving herd immunity. Since most of the leading vaccine supply was pre-ordered by wealthier nations, the nationalistic competition for vaccines is a major contributor to the inequitable distribution of the COVID-19 vaccine. Finally, this research concludes that the global community must take initiatives to enhance the rate of COVID-19 vaccinations in all countries, regardless of wealth [7]. While the findings in this study are significant, the context of vaccine usage is better represented by the indices presented in this study.

III. Data Collection and Cleaning

The data analyzed in this study include COVID-19 vaccination data and a broad range of country indicator statistics. The COVID-19 vaccination rate data used in this study were obtained from Our World In Data [8]. This expansive data set contains a range of COVID-related statistics for most countries, including cases, deaths, vaccinations, hospitalizations, tests, and a handful of relevant country indicators. Vaccine supply data were obtained from the Multilateral Leaders Task Force on COVID-19 [9], containing statistics for incoming, delivered and administered COVID-19 vaccines. Multiple data sets of country indicators were collected from the World Bank [10] and compiled into one centralized data set.

A. COVID-19 Vaccination Rate Data

This data set is structured as having one row per day and per country. These are presented chronologically within each country and by order of the country's ISO code. The starting day per country is variable when the country begins to submit COVID-19-related statistics, and the ending day is May 25, 2022. This date was chosen as most countries started vaccinating their population prior to May of 2021, aside from seven countries. One year is a sufficient period of time for countries to reach second-dose vaccinations. The raw data include the country's ISO code, continent, country name, date, and various COVID-19 statistics: cases, deaths, vaccinations, hospitalizations, tests, and some relevant country indicators. Due to the focus of this study being on vaccination rates, most of the non-vaccine-related data were discarded. The remaining vaccination data included the total number of vaccinations, people vaccinated, fully vaccinated, and the total number of boosters (count and per hundred people for each). Also remaining was new vaccinations (count, smoothed, smoothed per million) and new people vaccinated (smoothed, per hundred); however, these were discarded. Many countries in this set did not provide updated statistics every day. Dates that did not receive updated vaccination data were left with blank cell values. These empty cells were populated with the previously filled data from that country, resulting in a complete data set.

To effectively summarize this expansive vaccination data, a new table was created: COVID Metadata. This consisted of a country for each row and various statistics from the larger vaccination set. This data set contained a total of 208 countries and territories. The counts of people with at least one dose per hundred, people with two doses per hundred, and people with at least one booster per hundred were recorded for the extraction date: May 25, 2022. The first vaccination date was added, calculated as the first recorded day with vaccination counts from the COVID-19 data set. Lastly, current total populations were obtained from the COVID-19 data set. Table I displays these attributes.

TABLE I
INITIAL COLUMNS IN COVID METADATA DATA SET

COVID Metadata Columns
ISO Code
Continent
Country
People Vaccinated Per Hundred
People Fully Vaccinated Per Hundred
Total Boosters Per Hundred
Date of First Vaccination
Population

B. COVID-19 Vaccine Supply Data

Following the creation of COVID Metadata, supply data were obtained from the Multilateral Leaders Task Force on COVID-19 on May 25, 2022 [9]. This data set consists of a range of current supply data for each country, including expected, delivered, and administered vaccine counts. The column "Administered Doses (% of Delivered)" was the only vaccine variable selected from each country. Determining utilization of the COVID-19 vaccine is achieved by assessing how much of the vaccine supply available has been administered to the population. Countries that received more doses than required to doubly vaccinate their population will not achieve low utilization scores, as vaccination rates will be considered in this calculation. This column was merged into the COVID Metadata data set by the ISO Code for each country. 183 countries/territories were successfully merged between the two sets.

C. World Bank Data

Three data sets were collected from the World Bank's data catalogue: world development indicators [11], environment, social and governance data [12], and health nutrition and population statistics [13]. The world development indicators provide a basis for comparing countries' general quality of life. The environment, social, and governance data set and the health nutrition and population data both contain data that likely impact the vaccination rate.

Only data from 2019 were included for each of these data sets. At the time of data retrieval, 2019 data were determined to be the most up-to-date for the largest amount of countries while having validity as a recent year. The data were organized to reflect the following format: one row for each country and indicator values for each year as columns. This was done to align the data structure with the COVID Metadata data set previously created. A hindrance to the effectiveness of these data is a large number of countries and indicators with incomplete data. Specific countries do not supply data for many indicators, which proves an issue for analyzing data among these countries. To mitigate against this, indicators missing over a third of their values are ignored during the analysis. These resulting sets were combined with the COVID Metadata data set, using the country ISO code. This resulted in a complete representation of each country, including important vaccination statistics and collected indicators from the previous decade.

IV. METHODOLOGY

Multiple tools were utilized within RStudio [14] for this study. The VUI and VAI were first created from the data collected, using merging, scaling and equal-depth binning processes. Following this, the Pearson correlation method IV-B was applied to discover correlations between vaccination statistics, VUI/VAI scores, and country indicators within all data and subsets split on GDP per capita and population indicators. Throughout this process, data were manipulated, filtered, and visualized using the R language libraries in RStudio.

A. RStudio

RStudio was used to manipulate, filter, visualize, and analyze data. It is an open-source software used for data science [14]. It provides an interactive development environment for the data science language R. R contains many functionalities to analyze data. For example, filtering methods in R were used to derive countries with specific indicator values. The library ggplot2 [15] was used to visualize data in different ways, including histograms, scatter plots, correlation plots, and cluster plots.

B. Pearson Correlations

The Pearson [16] correlation coefficient was used to investigate the correlation between vaccine data and indicators. This statistical method is widely used to measure the degree of association between two continuous variables. This is represented in a range from +1 to -1. The two ends are perfect associations, with +1 as a positive and -1 as a negative. Only correlations higher than 0.500 or lower than -0.500 were recorded in this study.

V. INDICES

To best represent COVID-19 vaccine utilization and acceleration, VUI and VAI were created based on the collected data. Both indices score each country on a scale from 0 to 100, with the worst-performing country scoring 0 and the best scoring 100. Vaccinations per hundred, supply counts, date of first vaccination, and population variables are all used in either one or both indices. Table II displays each variable used in the calculation of one or both indices.

The DFV was modified to best represent the time between each country's first vaccination. The first vaccination (Latvia on December 5, 2020) was set to 1, with each consecutive day adding 1 to this count with the highest value of 321. For example, a country with its first vaccination on January

TABLE II
VARIABLES USED TO CALCULATE THE VUI AND VAI

Variable	Description
P_1D	% population with one dose at 150 days
P_2D	% population with two doses at 05/25/2022
PB	% population with a booster at 05/25/2022
DFV	Date of first vaccination (scaled)
AD	Administered doses (% of delivered)
POP	Total population (binned)

5, 2021, would have a DFV score of 31. Following this conversion, DFV values were scaled to values between 0 and 10, where the first country to vaccinate has a DFV of 0 and the last one scores 10. The following formula was used to scale these values:

$$f(x) = \frac{(a-b)(x-min)}{(max-min)} + b$$

In this formula, x is the value being scaled. Max and min represent the current maximum and minimum values of x. The variables a and b are the new maximum and minimum values to which x is to be scaled. With the previous minimum and maximum values being 1 and 321 and the new minimum and maximum values set as 0 and 10, the following formula exemplifies this scaling for Afghanistan:

$$DFV = \frac{(10-0)(86-1)}{(321-1)} + 0 = 2.66$$

The POP value was binned into ten equal-depth bins, with the remaining countries in an eleventh bin. The lowest 18 countries/territories were put in bin one, and the USA, China and India are in bin eleven. Binning was chosen over scaling these values due to China and India's substantially large populations, where scaling would reduce most countries' scaled values to be very low.

A. Vaccine Utilization Index

The VUI quantifies the utilization of the COVID-19 vaccine by comparing supply usage and vaccination counts. Both of these statistics are considered for a fair evaluation, and bias is added to the final score based on the DFV. The VUI formula is as follows:

$$\textbf{VUI} = \max\{(AD), (P_2D + (0.1 * PB))\} + DFV$$

The MAX function in this equation will take the maximum value between two values: AD and P_2D + (0.1*PB). AD is the ratio of administrated doses over received doses (total count of used and ready to be administered). This ratio demonstrates how much of a country's supply has been used on its population. On the other side, adding 0.1*PB to P_2D effectively represents how vaccinated populations are. Including 10% of the PB weight rewards countries for having high booster counts, rather than only vaccination counts represented by the P_2D.

Taking the maximum of AD and P_2D + (0.1*PB) rewards countries that have used most of their supply, vaccinated most of their population, or both. Many countries have not received enough doses to vaccinate their population fully but have used most of this supply. This indicates excellent utilization of what they have available, taken into account in the VUI. On the contrary, some countries have much more supply than required to doubly vaccinate each individual, and their AD will be low. In this scenario, if the country has a highly vaccinated population, it will still score well on the index. Countries that have neither used their supply well nor have high vaccination counts will not score well on this index. The last segment adds its scaled DFV values to this score. Adding this value allows countries that could not begin vaccinating earlier a better chance to score higher.

Adding a score to the VUI for higher DFV values results in some countries scoring higher than 100. To effectively compare each country's utilization against each other, the VUI was scaled to 0-100 with the same scaling formula used for the DFV. This does not change how each score relates to the other but serves as a better representation as a whole. The country to best utilize their vaccine is ranked as 100, and the country on the opposite end is ranked at 0. The mean VUI value out of 183 countries/territories calculated is 68.18, and the median is 74.67. Scores over 80 are adequate, using supply well and vaccinating most of the population. VUI values scoring over 90 indicate an excellent utilization of the COVID-19 vaccine supply.

Table VI illustrates the top and bottom 15 countries on the VUI scale. The top three VUI countries display incredible vaccination rates, and combined with the PB scores (and DFV for Brunei), these countries display excellent vaccine utilization. As observed in the bottom three ranked countries, poor vaccination rates and low AD scores result in poor VUI scores. Burundi displays significantly lower scores, with negligible vaccination rates and supply usage, with most of its points coming from the largest DFV value. However, after scaling the VUI, Burundi's lowest score resulted in a value of 0.

B. VAI

Rather than quantifying a country's overall vaccine utilization, the VAI represents the rate at which each country began to vaccinate during the first 150 days per country. 150 days were selected for ample time to ramp up vaccination efforts after the first dose in a country. The VAI formula is as follows:

$$\textbf{VAI} = P_1D + POP - DFV$$

Using P_1D rather than P_2D is ideal in this index due to efforts within the first 150-day vaccination period to vaccinate everyone once. Added to the P_1D value is the binned POP value. Higher POP results in a higher bonus for the VAI score, accounting for lengthier and more complex operations to vaccinate larger groups of people. The DFV was scaled just as done for the VUI: from 0 to 10, with the first country to vaccinate scoring 0. However, this DFV value is subtracted from the index total. Since the P_1D is calculated at each country's 150 days of vaccinating, countries that start later will benefit from adopting strategies developed by earlier countries and may have increased vaccine access as manufacturing

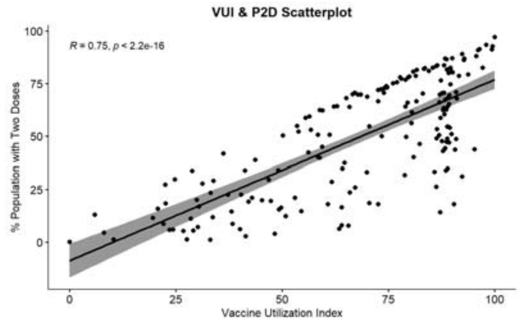

Fig. 1 Scatterplot of VUI and percent of the population with two doses

efforts increase. Therefore subtracting the DFV puts each country on an even playing field.

Keeping consistent with the VUI ranking, the VAI was scaled with the same formula to 0-100. Some values before this scale rank with negative scores, and scaling values from 0 to 100 represent better differences between the country's scores. The mean VAI value is 32.48, and the median is 26.61. Scores over 40 represent adequate efforts to vaccinate within the first 150 days, and any value over 60 indicates strong vaccination acceleration efforts.

Table VII illustrates the top and bottom 15 countries on the VAI scale. Cuba has the largest VAI by a wide margin due to its fast single dose vaccination, nearly 85% in only 150 days. As this scored the highest, it was scaled to 100. While Uruguay and Seychelles' P_1D scores are similar, the POP value added to Uruguay's final score results in a higher VAI than Seychelles'. Haiti's higher population does not make up for the late start, resulting in a subtraction of 1 between the two variables. Burundi had the least acceleration in its first 150 days of vaccinating, with the lowest P_1D score and the last to start vaccinating. It had a negative VAI score before scaling.

VI. Analysis

In this analysis, the VUI and VAI are compared with each other and the variables they represent. Next, sub-groups of countries are formed based on VUI/VAI, GDP per capita and population values. Pearson correlations are then generated between the indices and country indicators, identifying indicator types most correlated with VUI/VAI values.

A. Index Examination

These indices are examined in conjunction with other key vaccination variables to understand better the VUI and VAI. Fig.1 illustrates VUI values and their corresponding P_2D values for the same day used for VUI calculations (May 25, 2022). No P_2D is greater than its VUI value, as it would be taken as the maximum value in the formula. Where countries have utilized their supply to a higher ratio than $P_2D + (0.1 * PB)$, the VUI is greater than the P_2D value. Countries with low P_2D scores are, therefore, still able to achieve excellent VUI scores.

Fig. 2 illustrates VAI values with the same P_2D rates as Fig. 1. While P_2D is not used in the VAI calculation, a strong positive relationship exists between these two variables. Countries with high VAI scores also tend to have better vaccination rates. This indicates that countries with strong vaccination efforts at the start of their vaccination campaigns resulted in better rates by mid-2022. Some countries have high P_2D scores and very low VAI scores, unlike the previous countries, which indicates that some countries improved their practices as time proceeded in their campaigns.

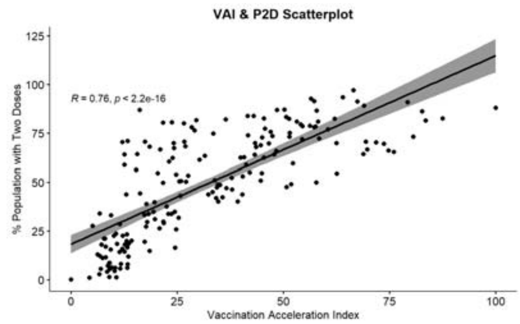

Fig. 2 Scatterplot of VAI and percent of the population with two doses

Fig. 3 plots both indices against each other in a scatter plot. A positive relationship is observed, and there is a large spread of VAI values at the higher VUI values. This finalizes the relationship between these indices; VAI indicates VUI rank among other countries. However, higher VUI scores have a larger spread of VAI. There is a large cluster of countries with low VAI scores and high VUI scores, depicting either low access to vaccine supply or a slow start that picked up pace as time proceeded. In a logical sense, a country's acceleration into vaccinating its population mostly indicates its vaccine supply and future vaccination rate utilization.

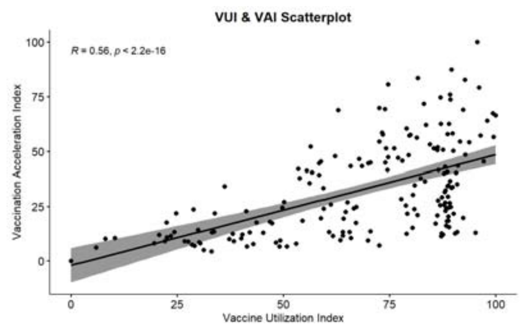

Fig. 3 Scatterplot of VUI and VAI

Figs. 4 and 5 display each index's relationship with the date of first vaccination. These plots contain the scaled DFV values, where 0 is the first country to vaccinate and ten is the last to start. Both indices typically observe higher index values with earlier first vaccination dates. The VUI is boosted with later DFV values to compensate for the lack of time each country has to use their supply and vaccinate their population. However, even with this boost, countries that started sooner report better utilization on average. The VAI and DFV relationship within the calculation is opposite of VUI, as later starting countries are reduced in VAI score to compensate for increased vaccination methodology development. This enhances the negative relationship between VAI and DFV values, where the higher the DFV (late start), the lower the VAI.

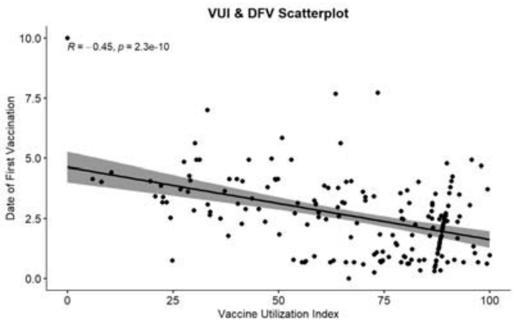

Fig. 4 Scatterplot of VUI and the date of first vaccination (scaled)

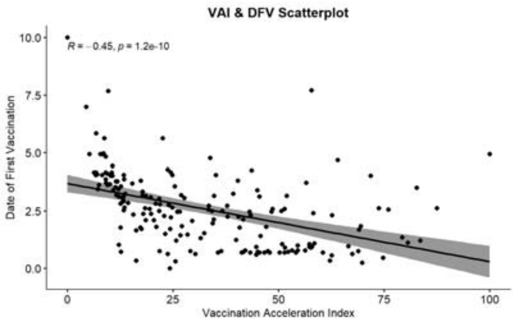

Fig. 5 Scatterplot of VAI and the date of first vaccination (scaled).

Figs. 6 and 7 illustrate a positive relationship between index rates and GDP per capita. This is more consistent in the VAI; however, all high-end GDP per capita countries have VUI scores higher than 80. VAI scores have higher variance than the line of best fit, but the line is more positive. Based on these two figures, wealthier countries could vaccinate more effectively within the first 150 days than less wealthy countries, on average. However, countries in the lower half of GDP per capita values had a wide range of vaccine utilization values, with most wealthy countries utilizing the vaccine well.

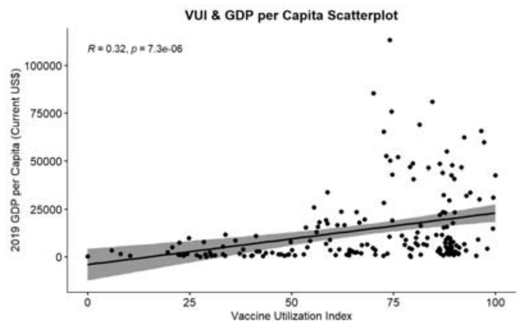

Fig. 6 Scatterplot of VUI and GDP per capita

Correlation tables between both indices and population indicators do not share as strong of positive or negative relationships observed in GDP per capita and DFV scatter plots. The spread of these countries along each index is large,

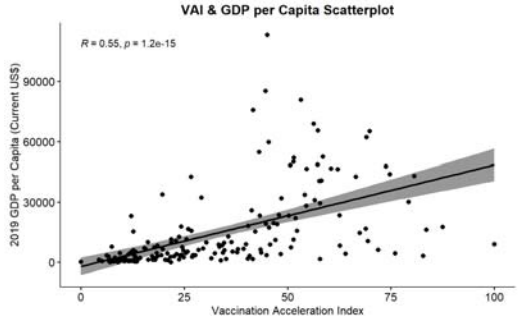

Fig. 7 Scatterplot of VAI and GDP per capita

and to effectively analyze these differences sub-groups must be made, isolating the extreme ends of this indicator.

Analyzing this data set alone does not tell the entire story of how other country indicators impact these indices. Isolating countries by the highest and lowest key values allows insight into what separates countries at these levels. The top and bottom 30 countries from two variables were split into four sub-groups, and in the following sections, these sub-groups are analyzed to the same degree as all data. These two variables are GDP per Capita (2019 data) and Population. These variables were chosen to give the increased context of how the counties on either extremity are separated in their VUI and VAI values. Choosing 30 countries per sub-group satisfies the following conditions: according to [17], and [18], 25 records are the minimum needed for effective Pearson correlations, and 29 is the minimum sample size to determine the reasonable high correlation between variables. Therefore, 30 is minimal enough to isolate index scores from the chosen variable (relevant for GDP per capita and population).

B. *Pearson Correlations*

With both indices discussed and sub-groups formed, the Pearson correlation method is applied to determine the highest correlations between indicators and index values. This process is completed with each data set, identifying the top most correlated indicators (higher than .500 or -.500) and their categories. These categories are defined as Series' by the World Bank [10]. Table III identifies all twelve indicator categories. Governance indicators in the correlation results are not classified in a series as they originate from the environment, social and governance datasets. This study will classify their series as Governance.

1) VUI Correlations: First, VUI Pearson correlations are generated for each data set, displayed in Table VI. Within All Data, negative correlations are observed for the mortality rate of infants and young children and the cause of death by communicable diseases and maternal, prenatal and nutrition conditions. Fig. 8 illustrates the negative correlation with neonatal mortality rate, displaying a larger cluster of countries with high VUI and low mortality rates. This assortment of mortality rate indicators is an interesting find; within all countries, those with higher mortality rates for young persons typically displayed lower vaccine utilization rates, suggesting that countries pre-disposed to younger deaths tended to also

TABLE III
INDICATOR CATEGORIES, ALL BUT LAST OBTAINED FROM THE WORLD BANK

Series
Economic Policy & Debt
Education
Environment
Financial Sector
Gender
Health
Infrastructure
Poverty
Private Sector & Trade
Public Sector
Social Protection & Labor
Social: Health
Governance

Governance indicators are defined in this study.

poorly utilize their vaccine supply. Also observed is a positive correlation with access to electricity, which reoccurs in other data set correlations within this analysis. This correlation is observed in Fig. 9.

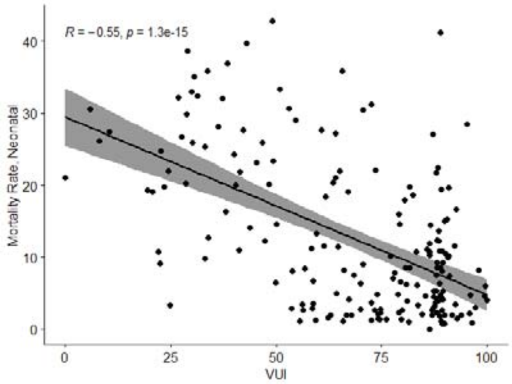

Fig. 8 Scatterplot of VUI and mortality rate, neonatal, with all countries

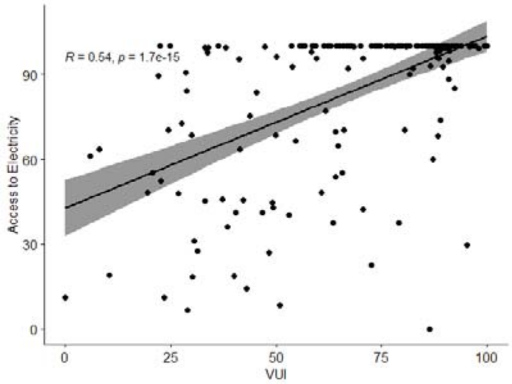

Fig. 9 Scatterplot of VUI and access to electricity, with all countries

Correlations between the VUI and GDP per capita comprise health factors, single governance, and environmental indicators. Within the bottom 30 GDP per capita countries, two of the highest correlations are with people using at least basic sanitation services in urban settings (presented in Fig. 10) and as a percent of the population. Negative correlations are observed with the age dependency ratio in two categories: young and percent of the working-age population. Based on this evidence, those lower-income countries that can provide their populations with these basic services and have (on average) fewer dependents per working-age individual were more effective in utilizing their COVID-19 vaccines. From the top 30 GDP per capita countries, only two indicators have Pearson correlations higher than (.500): immunization to BCG (% of 1-year-old children) and the voice and accountability estimate. While having high immunization rates for other diseases correlating with vaccine utilization is logical, the negative correlation with voice and accountability is more interesting. This suggests that within these 30 countries, those with less opportunity for the average person to have a political voice were more effective at utilizing their supply of vaccines. This may tie in with vaccine hesitancy, as those speaking against vaccine policies may have a smaller voice than other countries, minimizing the effectiveness of the anti-vaccination movement.

Significant correlations between VUI and population were only observed for the top 30 countries, with high correlations up to 0.766. Health indicators such as immunization to measles (% of children) and people using at least basic drinking water services have strong positive correlations, implying the importance of good health practices for countries to utilize the COVID-19 vaccine effectively. Fertility rate, birth rate, and population growth display negative correlations, suggesting that within these highly populated countries, those with slower and more controlled birth rates can also manage their vaccinations effectively. Access to electricity is yet again observed as a positive correlation. Immunization to measles and population growth correlations are observed in Figs. 11 and 12.

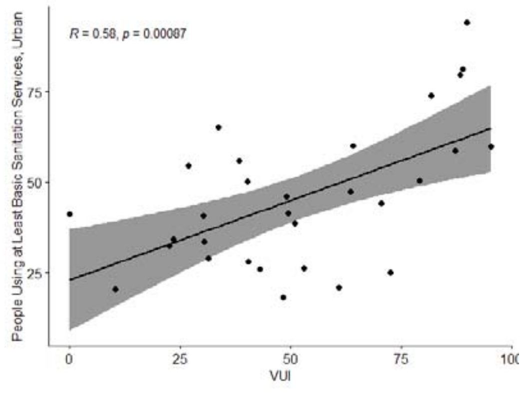

Fig. 10 Scatterplot of VUI and people that use at least basic sanitation in the urban population, with the bottom 30 GDP per capita countries

Comparing all significant VUI and country indicator correlations reveal a strong tie between health factors and COVID-19 vaccine supply utilization. Countries with better living conditions, such as lower mortality rates, access to electricity, and higher immunization to other diseases, display better rates of fully vaccinated individuals (and good supply usage if rates are low). Comparing the highest

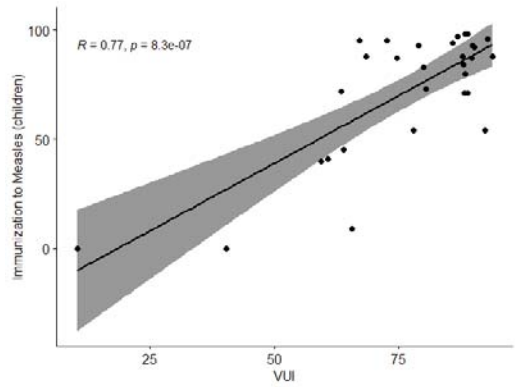

Fig. 11 Scatterplot of VUI and immunization to measles in children, with the top 30 populated countries

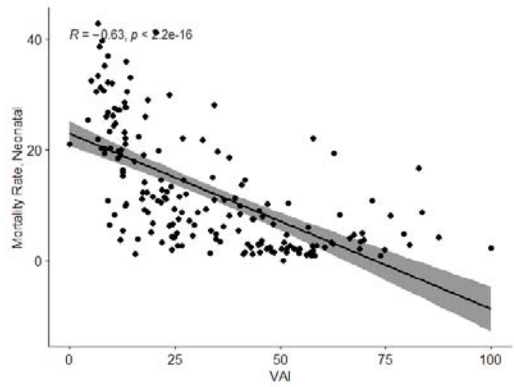

Fig. 13 Scatterplot of VAI and the neonatal mortality rate, with all countries

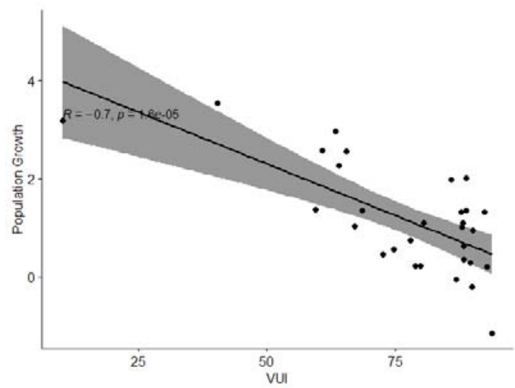

Fig. 12 Scatterplot of VUI and population growth, with the top 30 populated countries

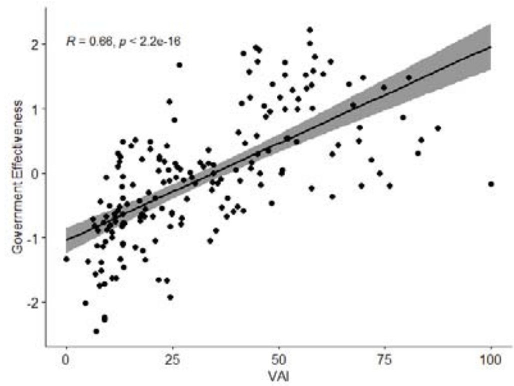

Fig. 14 Scatterplot of VAI and government effectiveness, with all countries

populated countries' correlations revealed that slower-growing populations also tend to yield stronger vaccine utilization.

2) VAI Correlations: Correlations between VAI and country indicator values have significant results. Table VII shows that mortality rate indicators negatively correlate between the VAI and all data, with the neonatal correlation displayed in Fig. 13. This is similar to what is observed between VUI and all data. Governance indicators such as government effectiveness (Fig. 14), the rule of law (Fig. 15), and control of corruption have high positive correlations suggesting that countries with well-structured governments could effectively roll out the vaccine in the early days of vaccination. Also observed is a negative correlation between birth rate and VAI, similarly associating slower population increase with index scores as discovered in the VUI correlations in top population countries.

While no indicators significantly correlate with the VAI in the top 30 GDP per capita countries, strong correlations are found within the bottom 30. Access to electricity and basic sanitation services are observed as high positive correlations, also present multiple times in VUI correlations. The access to electricity correlation is illustrated in Fig. 16. Survival of females to age 65 and life expectancy are both high positive correlations, enforcing the mortality rate of females as a negative one. Interestingly, the cause of death by communicable and non-communicable disease indicators directly oppose each other. Countries with high non-communicable disease causes of death also have good vaccine utilization (Fig. 17), while countries suffering from many communicable disease deaths have negative correlations. This follows the logical pattern that countries with lower rates of dying from contagious diseases would have better vaccine rollout procedures; however, only observing this difference in the bottom 30 GDP per capita countries is a notable find.

Access to electricity is the only notable correlation within the bottom 30 populated countries, with a 0.540 correlation. However, the strongest correlations observed yet are within the top 30 population countries and the VAI. Each of these correlations identified is valued between 0.798 and 0.901. The strongest are economic and governance indicators; regulatory quality, GDP per capita, GNI per capita, control of corruption, the rule of law, government effectiveness, the age dependency ratio (old), and voice and accountability. These high correlations speak volumes of the impact that a strong government had on early vaccine rates, especially when contrasting these highly populated countries. GDP per capita's correlation is exemplified in Fig. 18. These correlations also show a strong positive association between high counts of elderly individuals and VAI scores. While these correlations may be explained by global efforts to prioritize vaccinating elderly individuals first, this is only true if countries with higher elderly populations were given more vaccines per population than other countries. Another possible conclusion is that countries with higher elderly populations have increased

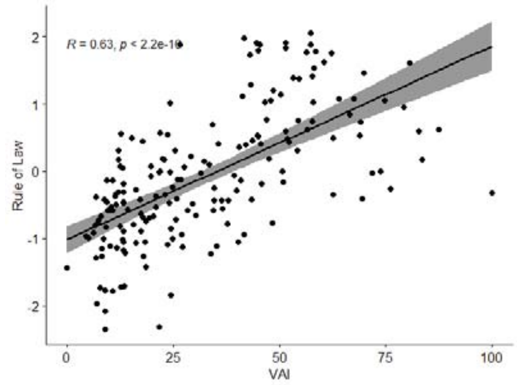

Fig. 15 Scatterplot of VAI and rule of law, with all countries

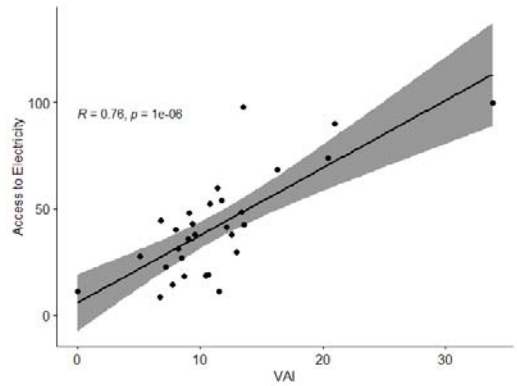

Fig. 16 Scatterplot of VAI and the rule of law, with the bottom 30 GDP per capita countries.

life expectancy, speaking for their overall quality of life. Based on these findings, the effectiveness of a country's government and wealth practices significantly influenced how effectively they could accelerate COVID-19 vaccinations within the first 150 days.

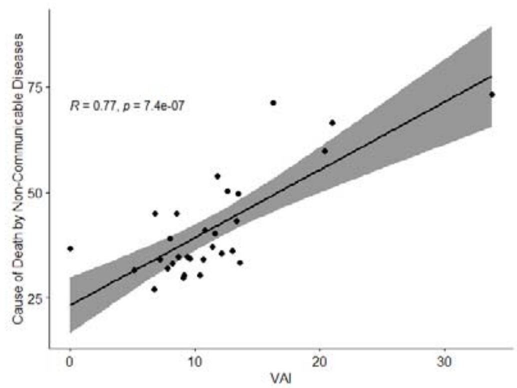

Fig. 17 Scatterplot of VAI and cause of death by non-communicable diseases, with the bottom 30 GDP per capita countries

Comparing each significant VAI and country indicator correlation reveals the significance of health and governance indicators and some economic and environmental indicators as observed in VUI correlations, health indicators such as mortality rates and people using basic sanitation services

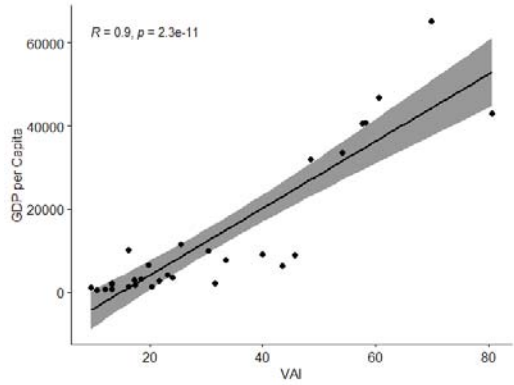

Fig. 18 Scatterplot of VAI and GDP per capita, with the top 30 populated countries

impact how effective a country was at vaccinating their population within the first 150 days of vaccinating. However, this index displays much stronger correlations with governance indicators, with strong correlations within the top 30 populated countries. The effectiveness of the government's ability to make laws regulate their people, control corruption, and economic standing all are highly correlated with how accelerated these vaccine rates were.

VII. Discussion

Contrasting VUI and VAI scores of each country revealed a cone effect from the lower to higher ends of both indices. Countries that did not effectively utilize their vaccine supply also struggled to vaccinate their population in the first 150 days of vaccinations. However, high VUI values matched a wide range of VAI values. Many countries with high early vaccination counts had low overall vaccinations and supply utilization by mid-2022, and just as many had positive scores on both indices. Comparing these scores with the date of first vaccination revealed that countries which started earlier scored higher on both fronts, regardless of the penalty for an early start in the VUI. GDP per capita displayed mostly positive correlations with both indices, while population values have insignificant correlations.

Examining Pearson correlations between country indicators and index scores displayed general trends across each front, with high correlations observed in specific sub-groups. Out of all VUI and country indicator correlations examined, the majority are categorized as Health, revealing that countries that tended to score higher on the VUI have better Health practices. This includes lower mortality rates and lower age dependency ratios. A higher rate of Governance and Economic indicators was observed within VAI and country indicator correlations, with just over half categorized as Health. The rule of law, regulatory quality and control of corruption are examples of these Governance indicators.

The Pearson correlations in this study depict VUI scores as heavily dependent on a country's overall Health. In contrast, countries leaned on effective government, economic well-being, and good health practices to score high VAI values. The negative correlation between mortality rates in both indices and the majority of correlations between health indicators and index scores, concludes that health practices are largely predictive of how effectively a country is able to vaccinate its population in both regards. Utility access, such as basic sanitation services and electricity, largely correlates with low-GDP per capita countries' index scores. The range of income groups observed in the top 30 populated highlights the impact that government effectiveness and economic status had on accelerating the vaccination rates in a country.

VIII. Recommendations

The lack of country indicator data is the biggest limitation of this study. Removing all indicators missing more than one-third of data across all countries reduced the indicator pool significantly. For example, no education indicators were able to be correlated with index scores. A smaller pool of countries that supply ample data would have to be used to analyze these indicators effectively. To be as inclusive as possible, these indicators were not analyzed. Another limitation is the reliability of COVID-19 vaccine rates data, as many countries do not update their data daily, resulting in vaccination rates that may not be accurate to the day.

A recommendation to further the findings of this study is to utilize the VUI and VAI. These indices account for two unique evaluations of countries' COVID-19 response effectiveness and can be used with other country evaluation data for further understanding.

IX. Conclusion

The VUI and VAI were created to summarize two key elements of countries' vaccination efforts: the utilization of vaccine supply and the acceleration of vaccination rates following the first vaccination in that country. Examining Pearson correlations between country indicators and these indices revealed what attributes of a country were associated with these two measures of vaccination performance. These findings conclude that the health quality of a country largely predicts overall COVID-19 vaccine utilization and early acceleration. This is observed to a higher degree in vaccine utilization; vaccination rates and supply usage directly reflect countries' value to their health excellence. While still impacted by health factors, the vaccination acceleration of countries within the first 150 days of vaccination was largely influenced by the effectiveness of its governing body. Governments with stronger law creation and enforcement practices, corruption mitigation policies, and economic status were more efficient in the early vaccination process than weaker governments. Other significant Pearson correlations within select sub-groups include higher index scores observed in high elderly populations, lower rates of population growth, and higher access to electricity.

Appendix

Tables IV and V identify the top 15 and bottom 15 countries in each of the two indices of this study.

TABLE IV
Top 15 and Bottom 15 Scoring Countries on the VUI

Country	VUI
United Arab Emirates	100.00
Brunei	99.50
Chile	99.36
Samoa	98.00
Denmark	97.22
Singapore	96.48
Malta	96.11
Cuba	95.74
Mozambique	95.26
Italy	93.79
South Korea	92.88
Bhutan	92.76
Qatar	92.39
South Africa	92.33
Belgium	91.64
Gabon	28.51
Yemen	27.59
Mali	26.77
Bulgaria	24.79
Senegal	24.39
Malawi	23.51
Togo	22.60
Saint Vincent & Grenadines	22.49
Syria	22.11
Namibia	20.78
Congo	19.58
Democratic Republic of Congo	10.38
Cameroon	8.06
Djibouti	5.90
Burundi	0.00

Tables VI and VII identify strongest Pearson correlations between VUI and VAI and country indicators within all data and each sub-group of countries.

TABLE V
Top 15 and Bottom 15 Scoring Countries on the VAI

Country	VAI
Cuba	100.00
Uruguay	87.49
Seychelles	83.55
Bhutan	82.74
United Kingdom	80.61
Malta	79.28
Mongolia	76.03
Israel	74.75
San Marino	73.73
Fiji	71.84
United States	69.80
Maldives	69.43
Qatar	69.01
Hungary	68.90
Chile	67.41
Madagascar	8.51
Congo	8.21
Guinea-Bissau	8.21
Benin	8.03
Central African Republic	7.79
Gabon	7.47
Sierra Leone	7.23
South Sudan	7.02
Lesotho	6.78
Chad	6.76
Papua New Guinea	6.67
Djibouti	6.25
Liberia	5.14
Haiti	4.39
Burundi	0.00

TABLE VI
CORRELATION OF VUI AND INDICATORS BY DIFFERENT DATA SETS

Data Set	Indicator	Correlation	Series
All Data	Cause of death, by communicable diseases and maternal, prenatal and nutrition conditions (% of total)	-.558	Health
All Data	Mortality rate, under-5, male (per 1,000)	-.549	Health
All Data	Mortality rate, neonatal (per 1,000 live births)	-.546	Health
All Data	Mortality rate, under-5 (per 1,000)	-.545	Health
All Data	Access to electricity (% of Population)	.544	Environment
All Data	Mortality rate, infant, male (per 1,000 live births)	-.543	Health
All Data	Mortality rate, under-5, female (per 1,000)	-.541	Health
All Data	Mortality rate, infant (per 1,000 live births)	-.540	Health
All Data	Mortality rate, infant, female (per 1,000 live births)	-.536	Health
All Data	People using at least basic sanitation services (% of population)	.529	Health
GDPPC Bottom	People using at least basic sanitation services, urban (% of urban population)	.576	Health
GDPPC Bottom	Cause of death, by non-communicable diseases (% of total)	.548	Health
GDPPC Bottom	People using at least basic sanitation services (% of the population)	.525	Health
GDPPC Bottom	Age dependency ratio, young	-.520	Health
GDPPC Bottom	Age dependency ratio (% of working-age Population)	-.519	Health
GDPPC Bottom	Access to electricity (% of Population)	.502	Environment
GDPPC Top	Immunization, BCG (% of one-year-old children)	.561	Health
GDPPC Top	Voice and Accountability: Estimate	-.533	Governance
Population Top	Immunization, measles second dose(% of children by the nationally recommended age)	.766	Health
Population Top	Access to electricity (% of Population)	.764	Environment
Population Top	People using at least basic drinking water services (% of Population)	.758	Health
Population Top	Age dependency ratio, young	-.745	Health
Population Top	Fertility rate, total (births per woman)	-.745	Health
Population Top	Age dependency ratio (% of working-age Population)	-.742	Health
Population Top	People using at least basic sanitation services (% of the population)	.738	Health
Population Top	Birth rate, crude (per 1,000 people)	-.706	Health
Population Top	Population growth (annual %)	-.701	Health
Population Top	Cause of death, by communicable diseases and prenatal and nutrition conditions (% of total) maternal,	-.658	Health

TABLE VII
THE CORRELATION BETWEEN VAI AND INDICATORS BY DIFFERENT DATA SETS

Data Set	Indicator	Correlation	Series
All Data	Government Effectiveness: Estimate	.658	Governance
All Data	Birth rate, crude (per 1,000 people	-.637	Health
All Data	Age dependency ratio, young	-.632	Health
All Data	Rule of Law: Estimate	.631	Governance
All Data	Mortality rate, neonatal (per 1,000 live births)	-.630	Health
All Data	Mortality rate, infant, male (per 1,000 live births)	-.626	Health
All Data	Mortality rate, infant (per 1,000 live births)	-.624	Health
All Data	Mortality rate, infant, female (per 1,000 live births)	-.622	Health
All Data	Control of Corruption: Estimate	-.621	Governance
All Data	Mortality rate, under-5: male (per 1,000,	-.617	Health
GDPPC Bottom	Cause of death, by non-communicable diseases (% of total)	.768	Health
GDPPC Bottom	Access to electricity (% of population)	.761	Environment
GDPPC Bottom	Cause of death, by communicable diseases and maternal, prenatal and nutrition conditions (% of total)	.729	Health
GDPPC Bottom	People using at least basic services, urban (% of urban population) sanitation	.707	Health
GDPPC Bottom	People using at least basic services (% of population) sanitation	.684	Health
GDPPC Bottom	People using at least basic sanitation rural (% of rural population) services,	.668	Health
GDPPC Bottom	Survival to age 65 female (% of cohort) ,	.654	Health
GDPPC Bottom	Mortality rate, female (per 1,000 female adults) adult,	-.643	Health
GDPPC Bottom	Life expectancy at birth, total (years)	.639	Health
GDPPC Bottom	Life expectancy at birth, female (years)	.636	Health
Population Bottom	Access to electricity (% of population)	.540	Environment
Population Top	Regulatory Quality: Estimate	.901	Governance
Population top	GDP per capita (current US$)	.896	Economic
Population top	GNI per capita, Atlas method (current US$)	.895	Economic
Population Top	Control of Corruption: Estimate	.879	Governance
Population Top	Rule of Law: Estimate	.865	Governance
Population Top	Government Effectiveness: Estimate	.849	Governance
Population Top	Population ages 65 and above (% of total population)	.841	Health
Population Top	Age dependency ratio, old	.827	Health
Population Top	Treatment for hypertension, male (% of male adults ages 30-79 with hypertension)	.806	Health
Population Top	Voice and Accountability: Estimate	.798	Governance

Acknowledgment

This research is supported by RiipenLevel Up Project Funding and MacEwan Undergraduate Student Research Initiative Project Funding.

References

[1] W. H. Organization. "Coronavirus disease (covid-19)." (), [Online]. Available: https://www.who.int/health-topics/coronavirus. (accessed: 03.15.2022).

[2] C. Wang, P. W. Horby, F. G. Hayden, and G. F. Gao, "A novel coronavirus outbreak of global health concern," *The Lancet*, vol. 395, pp. 470–473, 10223 2020. DOI: https://doi.org/10.1016/S0140-6736(20)30185-9.

[3] F. International Labour Organization, F. Agriculture Organization, and W. H. O. Agriculture Organization. "Impact of covid-19 on people's livelihoods, their health and our food systems." (2020), [Online]. Available: https://www.who.int/news/item/13-10-2020-impact-of-covid-19-on-people's-livelihoods-their-health-and-our-food-systems. (accessed: 03.15.2022).

[4] C. P. Bown and T. J. Bollyky, "How covid-19 vaccine supply chains emerged in the midst of a pandemic," *The World Economy*, vol. 45, pp. 468–522, 2 2022.

[5] N. E. MacDonald, "Vaccine hesitancy: Definition, scope and determinants," *Vaccine*, vol. 33, pp. 4161–4164, 34 2015. DOI: https://doi.org/10.1016/j.vaccine.2015.04.036.

[6] M. Sallam, "Covid-19 vaccine hesitancy worldwide: A concise systematic review of vaccine acceptance rates," *Vaccine*, vol. 9, pp. 468–522, 2 2021.

[7] P. Basak, T. Abir, A. Al Mamun, *et al.*, "A global study on the correlates of gross domestic product (gdp) and covid-19 vaccine distribution," *Vaccines*, vol. 10, no. 2, 2022.

[8] E. Mathieu, H. Ritchie, L. Rodés-Guirao, *et al.*, *Coronavirus pandemic (covid-19)*, https://ourworldindata.org/coronavirus, 2020.

[9] M. L. T. F. on COVID-19, *Vaccine supply & delivery*, 2022. [Online]. Available: https://www.covid19taskforce.com/en/programs/task-force-on-covid-19-vaccines, (accessed: 05.25.2022).

[10] W. Bank. "Data catalog." (2022), [Online]. Available: https://datacatalog.worldbank.org/home. (accessed: 03.15.2022).

[11] W. Bank, *World development indicators*, 2022. [Online]. Available: https://datacatalog.worldbank.org/search/dataset/0037712/World-Development-Indicators, (accessed: 02.01.2022).

[12] W. Bank, *Environment, social and governance data*, 2022. [Online]. Available: https://datacatalog.worldbank.org/search/dataset/0037651/Environment--Social-and-Governance-Data, (accessed: 02.01.2022).

[13] W. Bank, *Health nutrition and population statistics*, 2022. [Online]. Available: https://datacatalog.worldbank.org/search/dataset/0037652/Health-Nutrition-and-Population-Statistics, (accessed: 02.01.2022).

[14] RStudio Team, *Rstudio: Integrated development environment for r*, RStudio, PBC., Boston, MA, 2021. [Online]. Available: http://www.rstudio.com/.

[15] H. Wickham, *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016, ISBN: 978-3-319-24277-4. [Online]. Available: https://ggplot2.tidyverse.org.

[16] D. Freedman, R. Pisani, and R. Purves, *Statistics: Fourth International Student Edition* (International student edition). W.W. Norton & Company, 2007, ISBN: 9780393930436.

[17] F. N. David, *Tables of the Ordinates and Probability Integral of the Distribution of the Correlation Coefficient in Small Samples*. Cambridge University Press, 1938, ISBN: 9780521071697.

[18] M. A. Bujang and N. Baharum, "Sample size guideline for correlation analysis," *World*, vol. 3, no. 1, pp. 37–46, 2016.

Mohammed Elmorsy is an assistant professor at the Computer Science Department at MacEwan University. He obtained BSc (Honors, 2008) and MSc (in WSN security, 2011) in Computer and Systems Engineering, Zagazig University, Egypt. In 2017, he obtained PhD (in Path Intrusion Detection Reliability in Wireless Sensor Networks) in Computing Science, University of Alberta. His research interests include data mining, wireless sensor network reliability assessment; developing efficient routing protocols and network security mechanisms for wireless sensor networks;designing efficient resource allocation and management methodologies for IoT applications that guarantee certain quality of information (QoI) and quality of service (QoS) levels.

Mohamad El-Hajj is an associate professor at the Computer Science Department, at MacEwan University. He obtained a Master's degree in Computer Science at the Lebanese American University, Lebanon, in 1998. He obtained his PhD from the University of Alberta, Canada, in 2006. His PhD thesis work focused on finding frequent patterns in large datasets. Dr. El-Hajj's main interest is in applying data mining intelligent tools in real-life examples. He applies his research in different sectors such as medical, engineering and education.