Influences on Vaccination Rates for Covid-19

Valentina Miller Yasin Cifci Cecilie Kudlek

February 23, 2022

Abstract

The ongoing COVID-19 pandemic and accompanying regulations have impacted the lives of people globally. High vaccination coverage could slow the pandemic and mitigate its impact. Although vaccines are widely available in many countries, vaccination rates differ greatly across countries. Here, the influences of several socio-economic factors on vaccination rates are examined. The GDP per capita, the life expectancy and long working hours of employees have a major impact on vaccination rates.

1 Introduction

Since the first cases of COVID-19 appeared in Wuhan in December 2019, the disease has spread worldwide and become a global pandemic. More than 400 million infected people have been registered globally, including more than 5.8 million fatalities with a presumably high number of unreported cases [1].

To counteract the outbreak of the disease, regulations and lockdowns were imposed by many governments, which in turn imposed many restrictions on daily life. Despite this, the pandemic has not yet been stopped and the hopes of many people lie in vaccines, which were first developed at the end of 2020. Vaccination could slow down the infection, or at least weaken the symptoms of the disease, leading to fewer hospitalizations and thus keeping the health-care system intact. Many countries have launched their vaccination campaigns during the spring of 2021. At first, only vulnerable groups were able to get vaccinated, but toward the middle of the year, the vaccine was released to the entire population in most countries.

Although vaccines are promising, vaccination coverage varies widely across countries. The reasons for this have not yet been conclusively clarified. In many countries, vaccination is not mandatory, but the government encourages people to get vaccinated. Therefore, it could be that trust in the government has an impact on vaccination rates, along with other socio-economic factors.

The current study thus aims to investigate which socio-economic factors have an influence on vaccination rates for COVID-19.

2 Method

We started out the process by determining possible measures that we thought could have an influence on the vaccination rates. In the next step we started looking for sources for these measures and eliminated those for which there

weren't any available. The following is an overview of the datasets we ended up with as a result of this process.

2.1 Data

Since our primary concern is vaccination, it is crucial to have a good quality and reliable dataset on vaccination progress. Thankfully most countries provide frequent updates on these numbers and since they are provided by governments a certain level of quality is to be expected. Our dataset on vaccination progress was made available as a CSV file on Kaggle and originates from the "Our World in Data" Github-Page's "covid-19-data" dataset [2]. The total number of vaccinations for each country is collected by the OWID team from official government reports and is updated on a daily basis. We fixated the date for vaccination numbers to October 25th, 2021, because in November some countries started to announce that vaccination will get mandatory and we wanted to exclude this influence [3].

Our data on Covid deaths was made available on Kaggle and has its origin on Worldomoteter [4]. Worldometer collects and aggregates data from various sources on various topics and is used by governments such as the UK as well as cited by more than 16 thousand books and journals. The latest metrics on the Corona pandemic, including total Corona deaths, are updated around the clock by the Worldometer team and are again sourced from official government reports. The dataset we used was last updated on February 19th, 2021.

The data on secondary and tertiary education was collected by the Organisation for Economic Co-operation and Development(OECD). Specifically, it originates from a yearly publication named "Education at a Glance" [5]. It is giving an estimate for the probability of secondary and tertiary graduation rates for respective countries and was published on 2021-09-16. The survey was conducted in 2019.

Similarly, the data on trust in government originates from another annual publication of the OECD named "Government at a glance" which contains a survey regarding trust in government. The publication is from July 9th, 2021 and the survey was again conducted in 2019.

Unemployment data is provided by the World Bank. It was retrieved from the ILO database in December 2021 but the data was collected back in 2020. The ILO database is hosted by the International Labour Organization which operates under the United Nations.

GDP is also provided by the World Bank and is collected from multiple "EUROSTAT-OECD Methodological manual on purchasing power parities" publications. The given GDP values are the most recent values from each publication and the dates sometimes range back to 2011 but are mostly from 2020.

The Better Life Index was last published by the OECD in 2019 and is covering the year 2018. It is a single value quantifying the general well-being in a country and consists of 11 measures deemed by the OECD as essential.

After collecting the necessary data we performed pre-processing to get the data into a form we can work with. In the next steps we first extracted the relevant information on a per country basis. Then we transformed the data in such a way that we ended up with a single datapoint per country with a column for each of our measures. Finally we merged the data. It is important to mention that not all datasets contain data for the same countries so it was

necessary to perform an outer-join while doing the merging to not end up with only the countries contained in all datasets (which would be only 14).

2.2 Analysis

In the second part we visualized and analyzed that data. This needed to be done in such a way that the underlying relationships between the variables would be observable. For that we plot each of our measures with our target variable, which is the vaccination rate, on a two dimensional scatter plot that has the respective measure on the x-axis and the vaccination rate on the y-axis. This lets us visualize the relationship of each measure with the target variable. One important indicator of that relationship is the correlation coefficient. We expect some degree of linearity between a factor and the vaccination rate if that factor has an influence on the vaccination process. For that reason the correlation coefficient is a good measure to quantify the relationship. Additionally we visualize the correlation by fitting a linear regression line on the points of each plot. That way a high degree of correlation will be easily observable by a high concentration of points close to the regression line.

In the last step we calculated a p-value for each of the plots. The p-value indicates the probability of two random variables producing a dataset with the given correlation value. Given a dataset it gives us the probability of that data being produced by two independent random variables. In general a p-value of 0.05 and below is considered to be statistically significant. For that reason we only considered measures that had a value in that range to be influencing the vaccination rates.

We performed the analysis using a Python based stack including the packages Jupyter-Lab, Matplotlib, Sklearn, Scipy, Pandas and Seaborn. Our code can be found at `https://git.tu-berlin.de/valentina_miller2000/influences-on-the-covid-19-vaccination-rates/-/blob/master/`.

3 Results

Figure 1 reveals a high positive linear correlation between the share of people fully vaccinated and the gross domestic product per capita with a correlation coefficient of 0.701 and p-value of $7.54 \cdot 10^{-15}$, meaning it is very likely that these two variables are correlated. The same is shown by Figure 2 for the relationship between the share of people fully vaccinated and the life expectancy. These two variables have an even higher correlation coefficient of 0.782 and also a small p-value of $1.29 \cdot 10^{-07}$.

Figure 3 shows a negative linear correlation between the share of people fully vaccinated and the working hours of employees with a correlation coefficient of -0.48. The correlation is slightly weaker than for the previous two factors, but the small p-value of 0.006 indicates that correlation is very likely.

For the following factors, shown by figures 4-9, the p-value was always higher than 5%, which makes a correlation between the share of people fully vaccinated and the respective factors very unlikely.

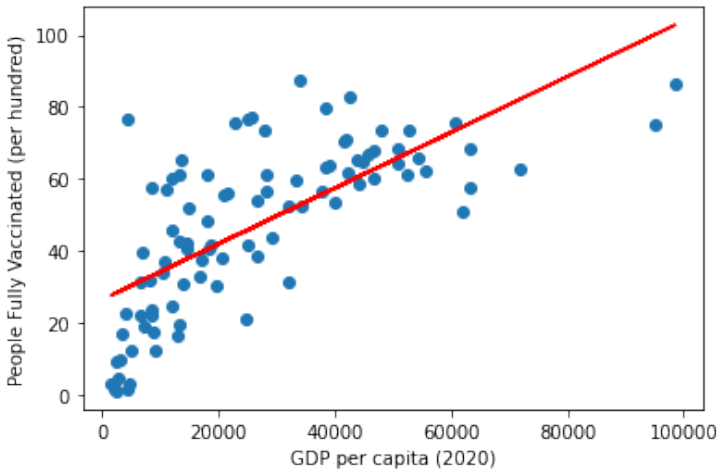

Figure 1: Correlation between share of population vaccinated against COVID-19 and GDP per capita

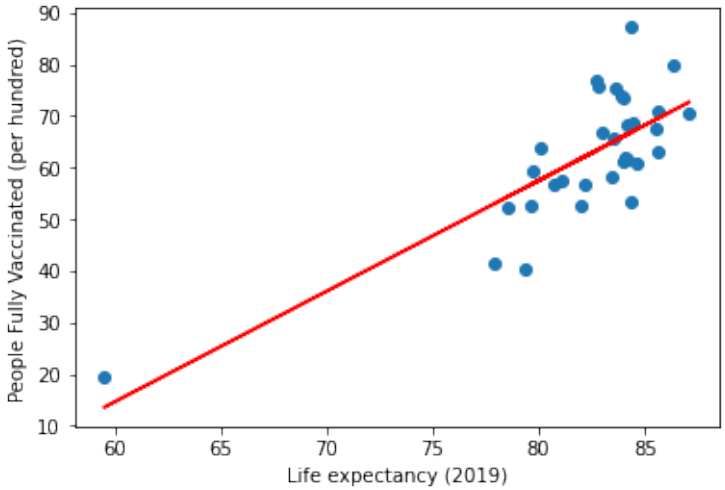

Figure 2: Correlation between share of population vaccinated against COVID-19 and life expectancy

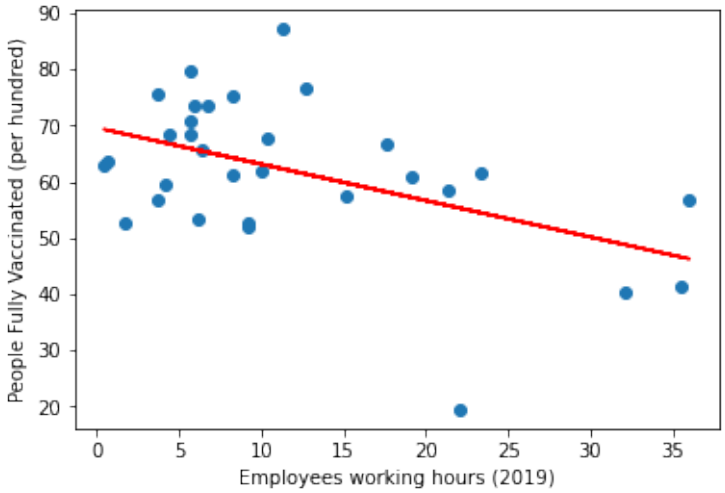

Figure 3: Correlation between share of population vaccinated against COVID-19 and employees working hours

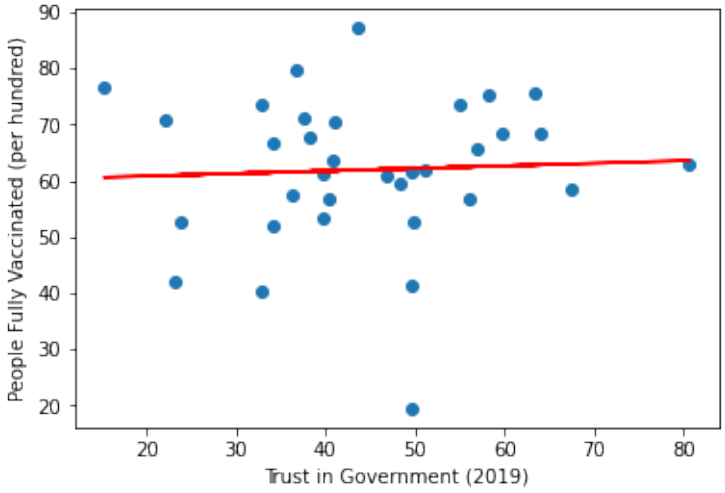

Figure 4: Correlation between share of population vaccinated against COVID-19 and trust in government

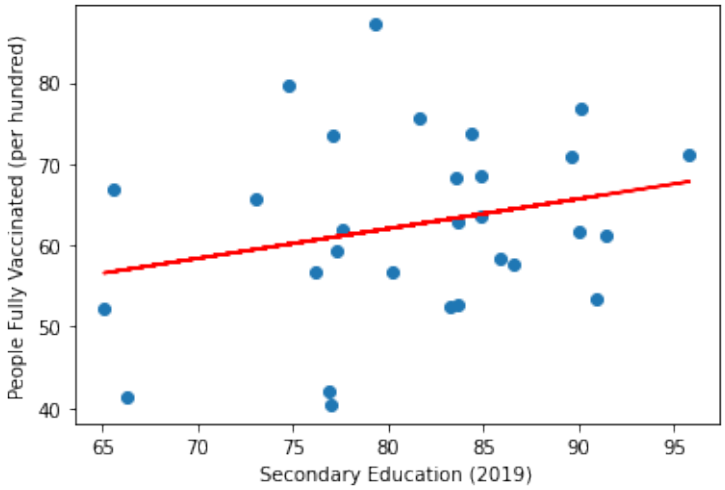

Figure 5: Correlation between share of population vaccinated against COVID-19 and secondary education

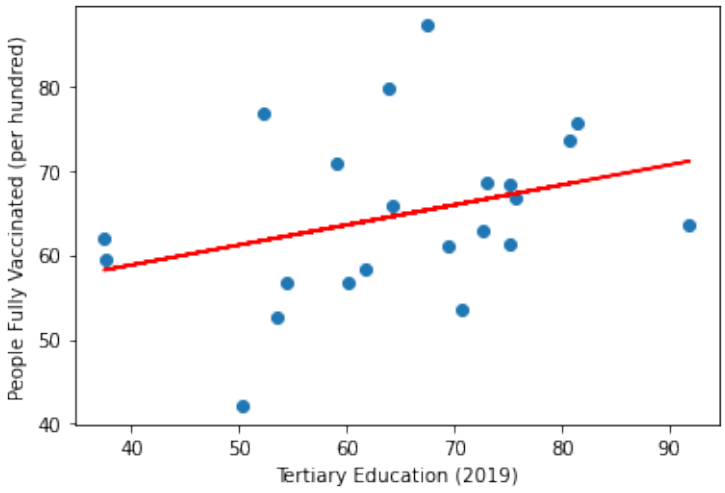

Figure 6: Correlation between share of population vaccinated against COVID-19 and tertiary education

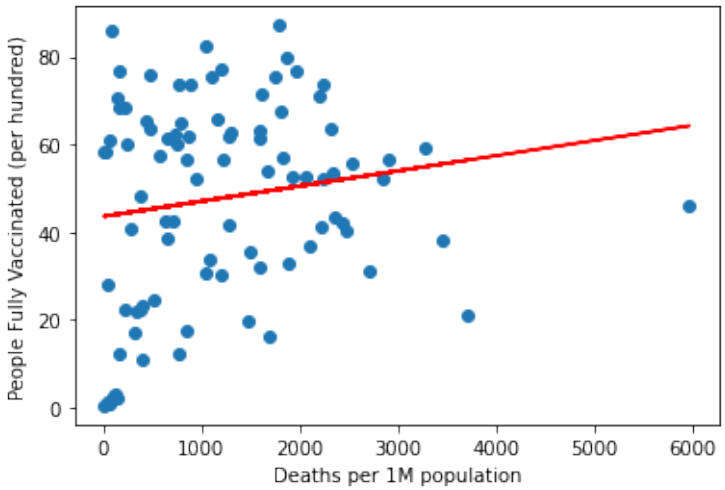

Figure 7: Correlation between share of population vaccinated against COVID-19 and deaths per 1 Million people

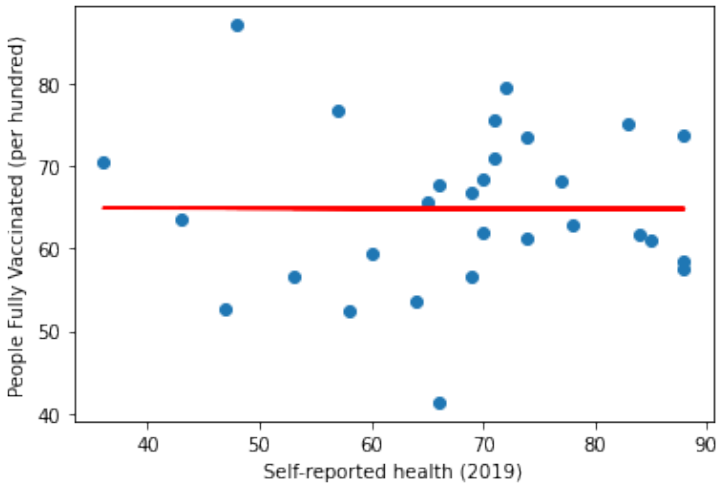

Figure 8: Correlation between share of population vaccinated against COVID-19 and self-reported health

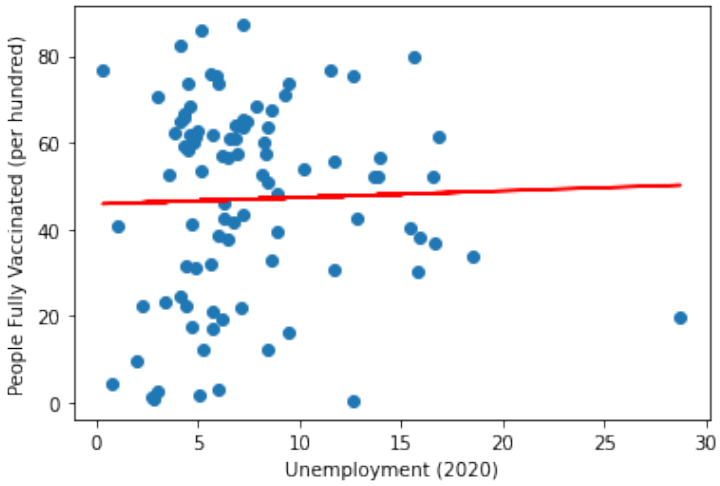

Figure 9: Correlation between share of population vaccinated against COVID-19 and unemployment

4 Discussion

The ongoing COVID-19 pandemic has cost the lives of many people and also damaged the economies of many countries. In addition to strict regulations by governments, which in turn restrict daily life, vaccination is one way to mitigate the effects of the pandemic, making a high vaccination rate desirable.

The current investigation therefore examined which factors have an influence on vaccination rates. The results show that there are three factors that have a relevant influence on the vaccination rate: the GDP per capita, the life expectancy as well as the percentage of people working more than 45hours per week. The higher the GDP per capita and the life expectancy the higher is the vaccination rate. However, the more people work long working hours in a country the lower is the vaccination rate of a country. The other factors did not have a statistically relevant influence.

The limitations of this study were a limited number of data points due to a limited number of countries. Moreover, data on socio-economic factors is hard to quantify, because some of the indices were collected by asking people about their attitudes and these findings are not entirely reliable.

It is important to mention that the GDP per capita and the life expectancy have a high relation among each other, they represent wealth. This is reasonable because a lot of poor countries have limited access to vaccines and therefore not even the possibility to get vaccinated even if they would want to. The more money a country has, the more it can spend on its health system. It was also shown that countries in which employees work longer have a lower vaccination rate. This could be because in poorer countries people earn less money and therefore have to work more to make a living. That's why another valuable investigation would be to look into the influences of vaccination rates for groups of countries separately, for example, to examine which socio-economic factors are of relevance within a group of richer countries with a high GDP versus within

a group of poorer countries with a lower GDP.

References

[1] World Health Organization. WHO Coronavirus (COVID-19) Dashboard. Accessed 2022-02-18. [Online]. Available: https://covid19.who.int

[2] Our World in Data. Data on COVID-19. Accessed 2022-02-18. [Online]. Available: https://github.com/owid/covid-19-data/tree/master/public/data#data-on-covid-19-coronavirus-by-our-world-in-data

[3] The Guardian. Austria plans compulsory Covid vaccination for all. Accessed 2022-02-18. [Online]. Available: https://www.theguardian.com/world/2021/nov/19/austria-plans-compulsory-covid-vaccination-for-all

[4] Worldometer. Worldometer. Accessed 2022-02-18. [Online]. Available: https://www.worldometers.info/coronavirus/#countries

[5] OECD, *Education at a Glance 2021*, 2021. [Online]. Available: https://www.oecd-ilibrary.org/content/publication/b35a14e5-en