An analysis of COVID–19 on the Dominican Republic and India

Emely Ortiz
Graduate School of Arts and Science
Fordham University
New York, United States
eortizsubero@fordham.edu

Titus Faroz
Graduate School of Arts and Science
Fordham University
New York, United States
tfaroz@fordham.edu

Rosa Sierra
Graduate School of Arts and Science
Fordham University
New York, United States
rsierra1@fordham.edu

Abstract – This paper presents a comparison between the COVID-19 positivity of the Dominican Republic vs India vs the World where we are assuming that the pandemic grew faster in Dominican Republic than in India, based on cases per million, and how the pandemic was developing throughout the time according to the responses of each government.

Keywords- Covid 19, India, Dominican Republic,

I. INTRODUCTION

A. Background information on COVID-19 on Dominican Republic:

The Dominican Republic is a small island with a population of 11 Million people, that every year puts 1.9% of its GDP towards healthcare, where the minimum established by the OPS is a 6%, cruel reality that came into the light when the pandemic was hit. Not only the amount of money that is going towards healthcare is badly invested, the dominican republic wasn't prepared economically, and neither educated enough epidemically speaking to understand the severity and complexity of the COVID-19, due to the inefficiency of the healthcare system to counterattack the COVID-19 pandemic the dominican Republic had to close from march until the beginning of the 2022, and made an investment of $800 million USD dollars just to fight the health necessities of the country.

Before the pandemic started the Dominican preparedness according to the GHS Index assesses countries is ranked 14th out of 34 countries in the Americas, with an overall score of 38, meaning the country's preparedness to respond to an epidemic is arguably low in absolute terms where the weakest areas refer to the absence of a plan to address human resources shortages, lack of evidence of an effective communication system with the population and between the public and private sectors in the event of a health emergency, low priority given to health worker care, and lack of a monitoring and evaluation plan.

An as result of this, the health system was in the process of generating and strengthening the primary care system, but it was facing the disarticulation of care, and an ongoing problem of lack of epidemiological plans, which lead to an outstanding amount of 620,237 infections and 4,833 dead in the span of 2 years, which represents a 6% of the population of the DR.

The positivity scale exponentially, and the spreads grew more rapidly than the response of the government, which led to a crash of the health system, with the bed occupancy of the Intensive care unit in average during the first 18 months in 79%, for every 1000 people there are 1.6 beds, when the recommended amount by the OMS is 4 for every 1000 people, 82% of public respirators in use and 42% of the private ones. But with a poor hospital infrastructure, mismanagement of funds, lack of equipment, and especially lack of beds. The hospitals were permanently overloaded and forced by circumstances to become COVID hospitals.

B. Background information on COVID-19 on India:

Global pandemic like covid 19 had a huge impact on India. India recorded the second highest number of confirmed cases in the world with 44,6M and third highest in number of deaths with 530,624. First case of covid 19 was reported on 30 January 2020 in Kerala and 43.8M cases recovered. India was imposed under lockdown on 25 March 2020. After the first outbreak of the disease the second one was more devastating with a shortage of vaccines, hospital beds, oxygen cylinders and different other medical supplies. With 3% of its gpd going towards healthcare . The statistics for the hospital bed are 0.53 beds per 1000 people, which is half bed per 1000 people and worse. India having 1.393B population, was a huge problem to provide with health facilities, food and other essential products.

It was a tough time every person business entity as whole of India was under lockdown for the protection of the spread of virus, on 31 march 2022 it was officially declared the lift on the lockdown.Over all standing of covid 19 in India can be represented by total of 3 waves of rise in cases of covid 19 where the second wave had the most impact on the whole country.

India's own vaccine COVAXIN by Bharat Biotech and other vaccines was a huge support for the people and the country with 2,199M number of doses and 950,7M fully vaccinated people, which sums upto 69.9 % of the population. Which is a huge number considering the total population of India.

C. Motivation:

For our purposes we would analyze the COVID-19 confirmed new cases evolution for the Dominican Republic and India, we select these countries only because of our backgrounds, and for personal reasons to compare our home countries in the covid matter. The idea would be to compare each country by the rest of the world and between each other.

This paper is divided as following: on section II, we would explain the origin of the used data set, on section III we would discuss the previous work made with this data set, on section IV we would explain the steps made from the data preprocessing previous to our analysis presented on section V with a series of graphs that supports our statements, finally, on section VI we would wrap all of ours conclusions on these analysis.

II. DATA SOURCE DETAILS

The data used for this study was extracted from the COVID-19 Data Explorer developed by Our World In Data [2]. In the same sense, their data about confirmed cases and deaths is provided by the Johns Hopkins University on their COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE), which collects this data from a variety of sources, including government and international organizations.

For reference, the data used on this analysis was downloaded on November 10, 2022.

III. PREVIOUS WORKS ON THIS DATASET

This dataset was collected by Edouard Mathieu et. al [1] for the purposes of developing the COVID-19 Data Explorer for the Our World In Data Webpage. This data explorer is an interactive interface which covers the main indicators on covid, such as new cases, deaths, vaccinations and others that can be analyzed by country. The idea of this paper is to focus our attention on these two countries and provide some explanation on the obtained graphs, since the Covid Explorer made by [1] only shows us the graphs but doesn't give a detailed description of the meaning of what we are seeing.

IV. DATA PREPROCESSING

A. Variable Selection: Our main variable of interest is the daily new cases of COVID-19 by country, Dominican Republic and India, but as we are also making a comparison of this countries with the rest of the world we need to add a need group name "World" which will be representing the summary of the entire world on our data. The dataset already contains this group as a country, so we didn't need to make the summarization ourselves. Then, for the daily new cases, as we are comparing different countries that means different populations, the variable used will be the new daily cases per million. Finally, we will have:

$X = [Dominican\ Republic,\ India,\ World]$

$Y = new\_cases\_smoothed\_per\_million$

Our data set is now reduced to: *country*, *date* and *new cases per million* filtered by our countries in X.

B. Null Values: Rows with null values were removed from the dataset. Another approach would be to compare null values between countries and remove for all countries the dates the rows (compared by date) in which another country has a null value, for a more fair comparison. But, since the null rows were around 20 of 3022 (0.66%) we didn't consider it necessary to take these measures.

C. Date Filtering: For a fair comparison we would consider the data on the date range: 2020-03-06 to 2022-11-06 , in which all of our countries have avaliable data. At this point we have 2,927 rows

D. Outliers: To maintain the integrity of the data, we won't discard outliers on the first analysis. Another analysis can be done removing outliers to compare the differences in the results.

E. Data Description: For the statistical analysis, we divide the data by the countries and look up for infinite values or wrong values on the data, but the number seems logical for each country and the world.

At the end of the data preprocessing we are left with around 2,927 rows, and the days between 2020-03-06 and 2022-11-06 are 976 days multiplied by the number of countries (3), if we have complete data we would have around 2,928 rows. As we have 2,927 rows we can conclude that our data preprocessing maintains the quality of our data based on the quantity of data. If we divide the quantity of data by country, we would have: for India, 975 rows ; for dominican Republic, 976 rows and for the World, 976 rows, which support our statement. Then, we would proceed to our analysis with this prepared data.

V. RESULTS AND INTERPRETATION

A. Histograms

For the Dominican Republic and India we see a left handed histogram distribution that means that the occurrence of a high indicator of cases per million was a rare occurrence in both countries. By this first view, we are having a sense that India had fewer cases of covid compared to the entire world and Dominican Republic, we would hold this hypothesis here for later testing and analysis with the rest of the graphs.
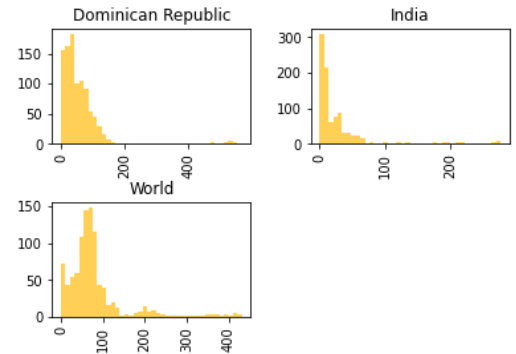

Figure 1. Histograms by country

B. Time Series Plot

The time series plot shows the evolution of the cases on the selected range, March 2020 to November 2022. In the case of the Dominican Republic, we could see only one higher wave, in the beginnings of 2022, that follows the same wage that have the entire world on this time, but the Dominican Republic wave was higher than the rest of the world and India, we can explain the occurrence of this wave by the december holidays, and the lack of measures in the Dominican Republic by that time.
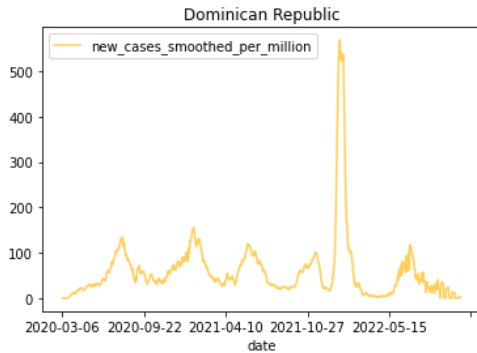

Figure 2. Dominican Republic Time Series

In the case of India, we see another higher wave, between April to June 2021,that can be explained as a result of the government lifting of restrictions, and the appearance of the delta variant, also the lax in taking precautions and negligence.
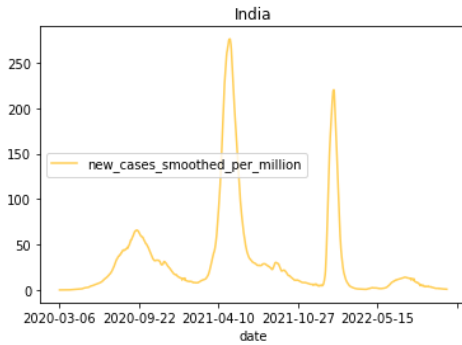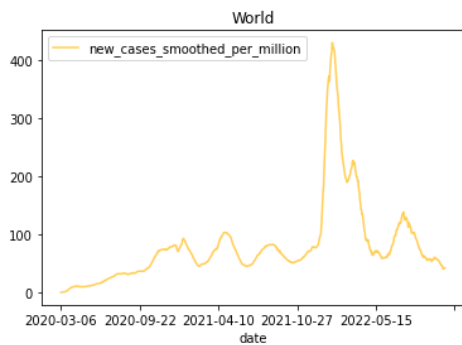

Figure 3. India Time Series

Figure 4. World Time Series

C. Cumulative plots of Y for X

The cumulative plot on this means the rate of growth of our new cases by time. For the Dominican Republic it maintains a steady growth line until January 2022, as seen on the time series plot this was the time of the highest wave of covid in the country, after this wage, we see the curve begins to flat.
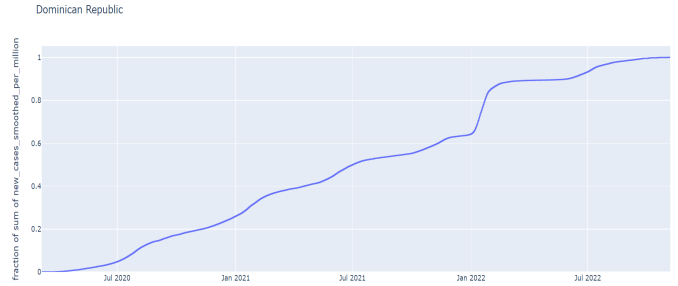

Figure 5. Dominican Republic cumulative plot

For India, the cumulative plot explains the three waves, but also remarks that outside these waves the covid situation in India remains with a little growth rate, and the curve flattens after every wave, as contrary to the Dominican Republic in which the plot is almost always growing, more similar to the world behavior.
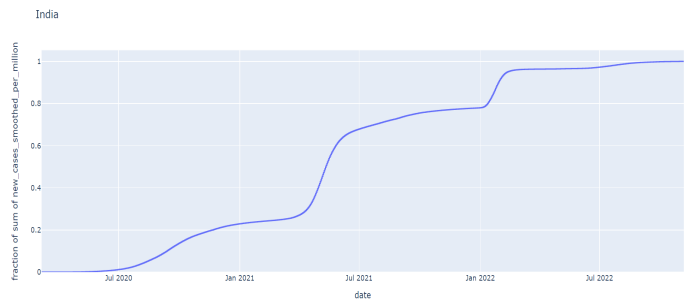

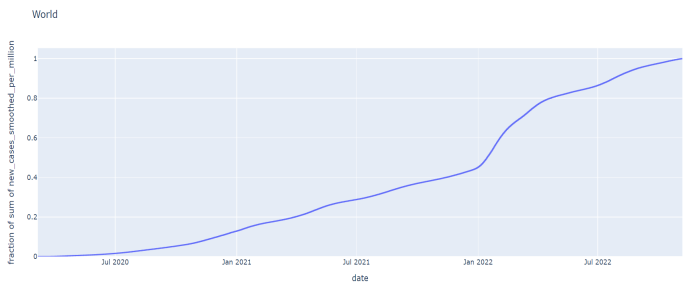**Figure 6.** India cumulative plot

Figure 7. World cumulative plot

D. Rate of Change of the new cases by country

For the rate of change calculation, the new cases per million was grouped by month, in the form of the mean, for each country. Then, we calculated the change between one month and the following month. By these graphs, we can see in general that India had a higher change between months, compared to the Dominican Republic, and in the same sense, the Dominican Republic had a higher change compared to the entire world.
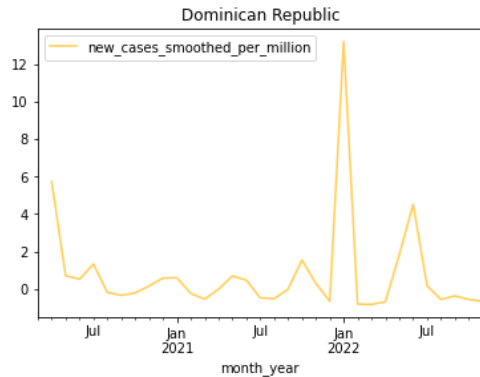

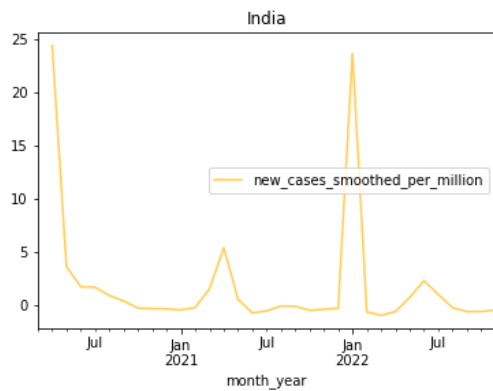
Figure 8. Dominican Republic rate of change

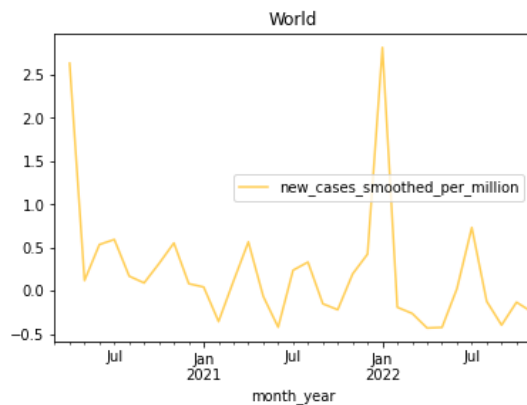**Figure 9.** India rate of change

Figure 10. World rate of change

E. Heatmap of Correlations for the new cases with other Attributes

The developed heatmap of correlations shows the relationship between the number of cases, deaths, vaccinations and covid test. For the Dominican Republic, on the figure 11, we see a medium correlation between the number of cases and the number of test, which is self explanatory, and a more weak correlation between the number of deaths and the new vaccinations, but since the vaccines weren't available since the beginning of the pandemic, we discard this correlation as a valid one.
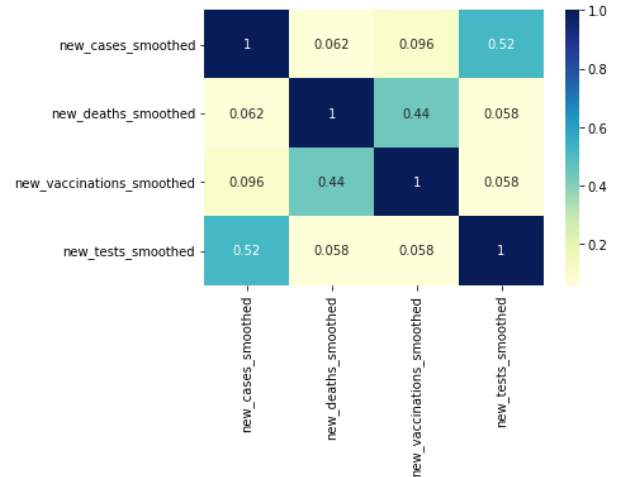

Figure 11. Dominican Republic heatmap of correlations

On the India heatmap of correlations, we see a strong correlation between the number of cases and the number of deaths, and for the number of covid test and number of deaths, this can be explained as more people dying can make other people nervous about the disease and get tested more frequently.
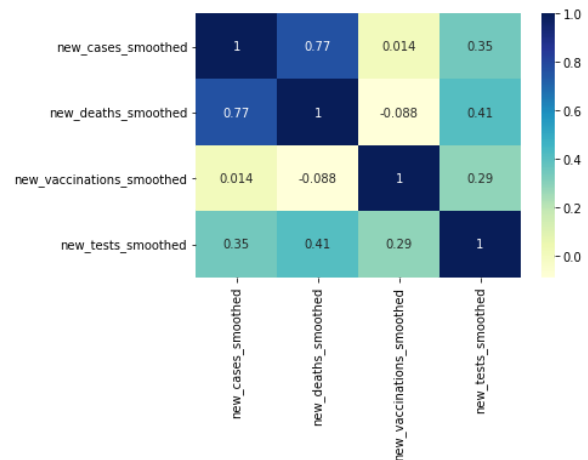

Figure 12. India heatmap of correlations

It's also important to remark that for either of the countries the new vaccinations had a relationship with the number of cases or deaths.

F. Barchart comparing Dominican Republic, India and the World for averages of new cases per million

The following bar chart compares the mean of new cases per million on the entire selected date range for Dominican Republic, India and the World. As we assume before this graph supports our hypothesis that the Dominican Republic had a higher number of cases compared by the population than India.
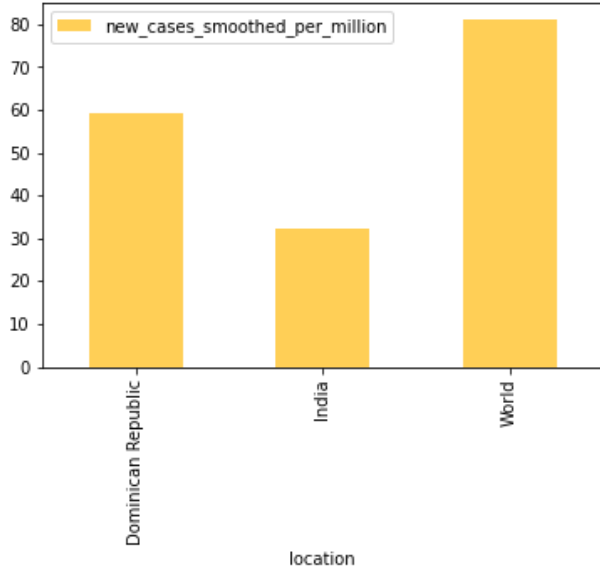

Figure 13. Barchart of average of new cases per country

G. Scatter Plots of HDI of All Countries Vs their current Y Colored by Regions (North America, South America, Europe, etc.)

The Human Development Index (HDI) represents the summary of statistics that consider the quality of life of a country, focusing on education, life expectancy and national income. The range of the HDI is between 0 and 1. The HDI of the Dominican Republic in 2019 was of 0.756 which is considered to be a high score, and the India was of 0.645 which is considered a medium score.

For our countries of interest, we would compare the number of new cases on the latest date avaliable: 2022-11-06, and the Human Development Index on the 2019, just before the pandemic hit. In figure 14, the Dominican Republic has higher HDI than India, and instead has a higher number of new cases per million than India.
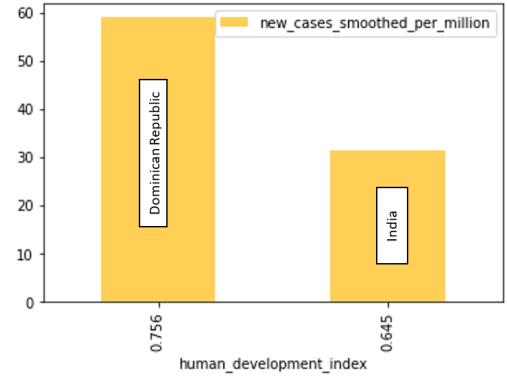

Figure 14. Barchart of new cases by HDI

On the figure 15, we see a summary by region of the number of cases per million at 2022-11-06 organized by the HDI.
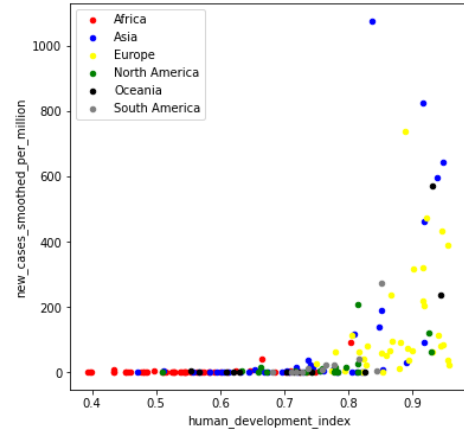

Figure 15. Scatter plot of new cases by HDI for region.

H. Test Hypothesis

Our idea with the hypothesis testing is to prove that the growing rates of covid for the Dominican Republic is higher than for India. Before testing the hypothesis, we would start with a most simple hypothesis comparing if the mean of the growing rates are equal for both countries, then if our hypothesis is rejected we would proceed to determine if the Dominican Republic is higher.

For the data given below, μ represent the average growthing covid cases by month, σ represents the standard deviation and n the number of rows, in this case the number of months.

$$\mu_{DR} = 0.766561$$
$$\sigma_{DR} = 2.680169$$
$$\mu_{INDIA} = 1.865128$$
$$\sigma_{INDIA} = 5.963937$$
$$n = 32$$

For our first hypothesis testing, we would analyze if the growth between the two countries are the same or if there is a significant difference.

$$H_0: \mu_{DR} = \mu_{INDIA}$$
$$H_1: \mu_{DR} \neq \mu_{INDIA}$$

Null Hypothesis (Ho): the mean of the growth was equally for both countries.

Alternative Hypothesis (H1): the mean of the growth was different between the countries.

$$significant\ level: 0.05$$

Now, we would calculate our number of degrees of freedom (df) given by the formula below:

$$df = \frac{\left(\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}\right)^2}{\frac{(s_1^2/N_1)^2}{N_1-1} + \frac{(s_2^2/N_2)^2}{N_2-1}}$$

$$df = 43.03 \equiv 40$$

Based on our significance level, 0.05, as this a two ends exercise, we would use 0.025, and the degrees of freedom, 40, we would calculate our comparative t-value:

$$t_{0.025,40} = 2.021$$

This t-value of 2.021 means the zone defined by $(-\infty, -2.021) + (2.021, \infty)$. If our calculated t-score is inside one of these zones, we need to reject our null hypothesis.

$$t\_score = \frac{\mu_{DR} - \mu_{INDIA}}{\sqrt{\frac{(\sigma_{DR})^2 + (\sigma_{INDIA})^2}{n}}} = -0.95043$$

Since our t-score of -0.95043 is outside the rejection zone based on our t-value, we need to accept the null hypothesis, then we can conclude that with a confidence level of 95% the mean of the growing rate of covid-19 cases for the Dominican Republic and India are equal. Then, our initial hypothesis that the Dominican Republic had a higher growing rate is automatically rejected, since the two countries were equal.

VI. CONCLUSION

The Covid-19 pandemic is a social and an economic crisis just as much as it is a health one – its repercussions, severe and far-reaching, were and are still being felt across the world.

It changed the way we communicate, care for others, educate our children, work and more. Over the past two years, the world has seen a shift in behaviors, the economy, medicine and beyond due to the COVID-19 pandemic.

As with any study, we used the best available data and understanding of mechanisms to develop possible inferences that make clear underlying decisions and uncertainty. And after applying the correct mechanism our initial hypothesis was rejected.

COVID-19 spread rapidly worldwide, but we have come to the conclusion that even with the different socioeconomic and cultural lifestyles that dominican, indians and the entire world had, the COVID-19 had a steady growth throughout the different phases that it had, with a decline after the completed vaccination phases, yes with variations over time from the different restrictions that went on in every country and how they were taken by the people.

REFERENCES

[1] Choudhury, S. R. (2021, April 30). India reports record Covid cases again — with over 386,000 new infections. CNBC. https://www.cnbc.com/2021/04/30/india-covid-crisis-more-than-6point6-million-cases-in-april.html

[2] Edouard Mathieu, Hannah Ritchie, Lucas Rodés-Guirao, Cameron Appel, Charlie Giattino, Joe Hasell, Bobbie Macdonald, Saloni Dattani, Diana Beltekian, Esteban Ortiz-Ospina and Max Roser (2020) - "Coronavirus Pandemic (COVID-19)". Published online at OurWorldInData.org. Retrieved from: 'https://ourworldindata.org/coronavirus' [Online Resource]

[3] Four Charts on the Dominican Republic's Crisis Response and Recovery. (2021, July 8). IMF. https://www.imf.org/en/News/Articles/2021/07/08/na071221-four-charts-on-the-dominican-republics-crisis-response-and-recovery

[4] García, D. G. (2022, May 27). Así luchó Dominicana contra el Covid con un sistema de salud deficiente. Bloomberg LÃNea. https://www.bloomberglinea.com/2022/05/24/asi-lucho-dominicana-contra-el-covid-con-un-sistema-de-salud-deficiente/

[5] Haczek, Á. R. (2021, May 28). Las UCI en el Gran Santo Domingo siguen saturadas; mayoría de contagios nuevos de covid-19 afectan la zona. CNN. https://cnnespanol.cnn.com/2021/05/27/santo-domingo-uci-covid-19-orix/

[6] N, O. (n.d.). covid-19-data/public/data at master · owid/covid-19-data.GitHub. https://github.com/owid/covid-19-data/tree/master/public/data.