Pontifícia Universidade Católica
do Rio de Janeiro

Jéssica Villar de Assumpção

Lessons learned from the COVID-19 pandemic in Latin America: a Data Science standpoint

Dissertação de Mestrado

Dissertation presented to the Programa de Pós–graduação em Engenharia de Produção of PUC-Rio in partial fulfillment of the requirements for the degree of Mestre em Engenharia de Produção.

Advisor : Prof. Paula Medina Maçaira Louro
Co-advisor: Prof. Fernanda Araujo Baião Amorim

Rio de Janeiro
September 2024

Pontifícia Universidade Católica do Rio de Janeiro

Jéssica Villar de Assumpção

Lessons learned from the COVID-19 pandemic in Latin America: a Data Science standpoint

Dissertation presented to the Programa de Pós–graduação em Engenharia de Produção of PUC-Rio in partial fulfillment of the requirements for the degree of Mestre em Engenharia de Produção. Approved by the Examination Committee:

Prof. Paula Medina Maçaira Louro
Advisor
Departamento de Engenharia Industrial – PUC-Rio

Prof. Fernanda Araujo Baião Amorim
Co-advisor
Departamento de Engenharia Industrial – PUC-Rio

Prof. Rafael Santos Erbisti
UFF

Prof. Leonardo dos Santos Lourenço Bastos
Departamento de Engenharia Industrial – PUC-Rio

Prof. Igor Tona Peres
Departamento de Engenharia Industrial – PUC-Rio

Rio de Janeiro, September 2nd, 2024

Jéssica Villar de Assumpção

Jéssica Villar is a senior data analyst at iFood and graduated from PUC-Rio with a bachelor's degree in Production Engineering. She was a researcher on the Effectiveness of Vaccination Against COVID-19 in Brazil using Mobile Data (EFFECT- Brazil). This is one of the 10 projects selected from more than 400 proposals for the International COVID-19 Data Alliance (ICODA) Data Science Initiative Grand Challenge and funded by the Gates Foundation. Her main areas of interest are Operations Research, Time Series, Machine Learning, Forecasting and Simulation Methods.

Bibliographic data

Villar de Assumpção, Jéssica

Lessons learned from the COVID-19 pandemic in Latin America: a Data Science standpoint / Jéssica Villar de Assumpção; advisor: Paula Medina Maçaira Louro; co-advisor: Fernanda Araujo Baião Amorim. – 2024.

44 f: il. color. ; 30 cm

Dissertação (mestrado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Industrial, 2024.

Inclui bibliografia

1. Engenharia de Produção – Teses. 2. COVID-19. 3. América Latina. 4. XGBoost. 5. SHAP Values. 6. medidas de contenção. I. Paula Medina Maçaira Louro. II. Fernanda Araujo Baião Amorim. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Industrial. IV. Título.

CDD: 004

To my family and friends
for their support and encouragement.

Acknowledgments

First of all, I want to thank my mother Cláudia, my aunt Ceiça and my grandmother Maria. Without their support, I wouldn't be where I am today. They are my inspiration and my strength.

I am very grateful to my advisors, Paula and Fernanda, who are much more than advisors to me, they are true partners. We have been together since graduation, and it has been incredible working alongside them.

I would like to thank health experts Ana Carla Pecego da Silva, Jorge Salluh, José Cerbino Neto, Fernando Bozza and Leonardo Bastos for their contributions with hypotheses and insights on this research.

I am very lucky to have the friends who have accompanied me throughout this journey. Thank you very much for always giving me strength Adailton, Vivian, Lucas and Bárbara.

I am very grateful to Ronaldo, my former manager at OLX, for his unconditional support during the period in which I had to balance work with my master's degree. Without his help, obtaining this master's degree might not have been possible.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

Abstract

Villar de Assumpção, Jéssica; Paula Medina Maçaira Louro (Advisor); Fernanda Araujo Baião Amorim (Co-Advisor). **Lessons learned from the COVID-19 pandemic in Latin America: a Data Science standpoint**. Rio de Janeiro, 2024. 44p. Dissertação de Mestrado – Departamento de Engenharia Industrial, Pontifícia Universidade Católica do Rio de Janeiro.

In the 21st century alone, the world has faced the devastating impacts of three acute respiratory diseases: Middle East Respiratory Syndrome (MERS), Severe Acute Respiratory Syndrome (SARS), and COVID-19, which evolved into a pandemic. These diseases have not only caused a large number of deaths but have also damaged the economies of the affected regions. In particular, countries in the Latin American and Caribbean (LAC) region have faced additional challenges due to greater social inequalities, limited access to health services, and precarious living conditions. Therefore, it is imperative to understand the effects of mitigation actions to guide actions to mitigate the health and socioeconomic impacts if (or when) new acute respiratory diseases emerge, especially in these countries. A retrospective study was conducted to model the dynamics of variation in COVID-19 mortality in LAC countries and analyze its association with vaccination strategies, containment measures, mobility restrictions, and socioeconomic factors. The study methodology applied clustering techniques that revealed two distinct clusters based on sociodemographic characteristics, followed by the application of XGBoost to model the dynamics of variation in deaths in the countries of each cluster, over time. Finally, the SHAP Values technique was applied to understand the associations between mortality and factors such as vaccination, containment measures and mobility restrictions. In addition, a panel of experts was held to assess the relevance and effectiveness of the results found. The study provides evidence that economic support and the completion of the vaccination scheme were especially relevant in reducing COVID-19 mortality. It was possible to detect two distinct groups of countries, where one group may have characteristics of greater vulnerability than the other group. The most important interventions for understanding COVID-19 mortality varied in two distinct periods of the pandemic: pre-vaccination and post-vaccination. In the pre-vaccination period, containment measures were the most important interventions for mortality in the least vulnerable countries, while for the most vulnerable countries, they were variations in population mobility. In the post-vaccination period, vaccination coverage was the most important intervention for mortality in the least vulnerable countries, while the most vulnerable countries were more impacted by containment measures.

Keywords

COVID-19; Latin America and the Caribbean; XGBoost; SHAP Values; containment measures.

Resumo

Villar de Assumpção, Jéssica; Paula Medina Maçaira Louro; Fernanda Araujo Baião Amorim. **Lições aprendidas com a pandemia de COVID-19 na América Latina: uma perspectiva de Ciência de Dados**. Rio de Janeiro, 2024. 44p. Dissertação de Mestrado – Departamento de Engenharia Industrial, Pontifícia Universidade Católica do Rio de Janeiro.

Somente no século XXI, o mundo enfrentou os impactos devastadores de três doenças respiratórias agudas: a Síndrome Respiratória do Oriente Médio (MERS), a Síndrome Respiratória Aguda Grave (SARS) e a COVID-19, que evoluiu para uma pandemia. Essas doenças não apenas causaram um grande número de mortes, mas também prejudicaram a economia das regiões afetadas. Em particular, os países da região da América Latina e Caribe (LAC) enfrentaram desafios adicionais, devido a maiores desigualdades sociais, acesso limitado a serviços de saúde e condições de vida precárias. Portanto, é imperativo compreender os efeitos das ações de mitigação para orientar as ações no sentido de mitigar os impactos sanitários e socioeconômicos, se (ou quando) surgirem novas doenças respiratórias agudas, especialmente nestes países. Foi realizado um estudo retrospectivo para modelar a dinâmica da variação da mortalidade por COVID-19 em países da LAC e analisar sua associação com estratégias de vacinação, medidas de contenção, restrições de mobilidade e fatores socioeconômicos. A metodologia do estudo aplicou técnicas de clustering que revelaram dois agrupamentos distintos com base em características sociodemográficas, seguidos pela aplicação do XGBoost para modelar a dinâmica de variação de mortes nos países de cada cluster, ao longo do tempo. Por fim, foi aplicada a técnica de SHAP Values para compreender as associações entre mortalidade e fatores como vacinação, medidas de contenção e restrições de mobilidade. Além disso, foi realizado um painel com especialistas para avaliar a relevância e efetividade dos resultados encontrados. O estudo fornece evidências de que o suporte econômico e a conclusão do esquema de vacinação foram especialmente relevantes para reduzir a mortalidade por COVID-19. Foi possível detectar dois grupos distintos de países, onde um grupo pode ter características de maior vulnerabilidade do que o outro grupo. As intervenções mais importantes para entender a mortalidade por COVID-19 variaram em dois períodos distintos da pandemia: pré-vacinação e pós-vacinação. No período pré-vacinação, as medidas de contenção foram as intervenções mais importantes para a mortalidade nos países menos vulneráveis, enquanto para os países mais vulneráveis, foram as variações na mobilidade populacional. No período pós-vacinação, a cobertura vacinal foi a intervenção mais importante para a mortalidade nos países menos vulneráveis, enquanto os países mais vulneráveis foram mais impactados pelas medidas de contenção.

Palavras-chave

COVID-19; América Latina; XGBoost; SHAP Values; medidas de contenção.

Table of contents

List of figures

List of tables

List of Abreviations

LAC – Latin America and the Caribbean

XGBoost – Extreme Gradient Boosting

SHAP – SHapley Additive exPlanations

DBSCAN - Density-Based Spatial Clustering of Applications with Noise

Those who pass by us, do not go alone, and do not leave us alone; they leave a bit of themselves, and take a little of us.

Antoine de Saint-Exupéry, *The Little Prince.*

1 Introduction

Since 2001, three new coronaviruses have been responsible for causing respiratory diseases whose spread led to four large-scale outbreaks (KWOK et al., 2019): the Severe Acute Respiratory Syndrome (SARS) in 2003, two Middle East Respiratory Syndrome (MERS) - the first in the Middle East in 2012 and the second in South Korea in 2015 -, and the COVID-19 pandemic that began in 2019. As the most recent outbreak, COVID-19 is the current case study for the ongoing challenge of emerging infectious pathogens (FAUCI; LANE; REDFIELD, 2020). Its causing virus, SARS-Cov-2, was first detected in Wuhan, China, in December 2019, and in February 2020, it had already spread to Latin America and the Caribbean (LAC) (BURKI, 2020). By March 2020, every country in LAC had reported infections (GARCIA et al., 2020), and by June 2020, LAC accounted for 27% of deaths from COVID-19 worldwide, becoming the region with the highest number of deaths in the world (PABLOS-MéNDEZ et al., 2020).

LAC is the world's most inequality-ridden region, with 53% of its working population earning their income from informal work (GARCIA et al., 2020). Underfunded state-run hospitals are the only source of medical care for those with informal jobs or unemployed (LITEWKA; HEITMAN, 2020). LAC countries suffer from social inequalities, several localities lacking access to healthcare services, and poor health outcomes. Overcrowding, limited sanitation, food insecurity, and poor nutrition are common health problems (LITEWKA; HEITMAN, 2020). Urban slums in large LAC cities, such as Bogota, Buenos Aires, Lima, Mexico City, Rio de Janeiro, and São Paulo, are more susceptible to COVID-19 and other infectious diseases due to their high population density (GARCIA et al., 2020). According to (PANIZ-MONDOLFI et al., 2020) and (PERES et al., 2021), the impacts differ for each country because of that heterogeneity. For example, the effect of COVID-19 will be more devastating in Venezuela than in more developed economies, such as Brazil, due to the Venezuelan humanitarian crisis, spreading many other diseases over the region (PANIZ-MONDOLFI et al., 2020).

Governments implemented several social distancing measures to cope with the COVID-19 pandemic (BARGAIN; AMINJONOV, 2021). Those policies were crucial to reduce the spread of the disease, especially before vaccines became available; however, they may also worsen the situation for the poor population, who lack food supplies, cannot work remotely, and rely on manual labor (BARGAIN; AMINJONOV, 2021). Lockdowns in Chile positively reduced the number of cases within high-income geographical areas, with no effects in lower-income areas (BENNETT, 2021). A possible explanation is that poor people are less likely to comply with social distance measures, as they must work (BARGAIN; AMINJONOV, 2021). Thus, mobility restrictions not accompanied by social transfer programs are less likely to be followed by the poorest population (BARGAIN; AMINJONOV, 2021; MARTINEZ-VALLE, 2021). In 2020, despite the demand from WHO that high equitable immunization coverage be prioritized at all levels (national, municipal, and district) (CHAN et al., 2022), the inequality in routine immunization in several LAC countries worsened with the scenario imposed by COVID-19, mainly due to the disproportionate impact on vulnerable populations (CHAN et al., 2022). A Brazilian study about inequity in COVID-19 vaccination showed that municipalities with low Human Development Index (HDI) had a lower first dose coverage than those with medium and high HDI (BASTOS et al., 2022).

Given these disparities and the complex interplay between socioeconomic factors and public health measures, it is important to deepen the understanding of the specific impacts and effectiveness of these interventions. Therefore, the objectives of this research are:

1. To understand which and how LAC countries are similar according to their socioeconomic and demographic characteristics;

2. To identify the best lag of days between interventions (vaccination coverage, containment measures, economic support, and population mobility) and COVID-19 mortality;

3. To identify which interventions are most important to understand COVID-19 mortality and how they are associated with COVID-19 mortality;

4. To propose public health guidelines for future respiratory diseases.

The research started by clustering countries with similar socioeconomic and demographic characteristics to achieve the proposed objectives. Next, the researchers separated into pre- and post-vaccination periods to model the relationship between variables and COVID-19 deaths in each cluster using XGBoost models. Finally, the association between each independent variable and mortality was interpreted by calculating its SHAP Values. Finally, the results were presented to health experts, who evaluated whether they were surprising, whether they contributed to creating public health guidelines, and what insights and bottlenecks they identified in the research.

This document is structured as follows. Chapter 2 presents some previous work relevant to the problem. In Chapter 3, the methodology are explained. In Chapter 4, the results are shown. Finally, Chapter 5 presents conclusion and future work.

2
Previous work

The objective of (OLIVEIRA et al., 2021) article is to investigate the demographic, clinical, and epidemiological factors associated with COVID-19 mortality in Rio de Janeiro, using data from the Unified Health System. The study employs advanced machine learning techniques, specifically the XGBoost model and the SHAP Values technique, to analyze the importance of several variables in predicting mortality among confirmed cases.

The steps of the study included collecting data from 243,509 confirmed cases of COVID-19, categorizing these cases into influenza-like illness (ILI) and severe acute respiratory syndrome (SARS), and analyzing these data through the XGBoost model. The model was used to develop robust predictions about patient outcomes. To interpret the model and understand the contribution of each variable in the predictions, SHAP Values were used, which provide a detailed view of the impact of each characteristic. The study results indicated that factors such as advanced age, especially over 60 years, black race, and the presence of comorbidities such as heart disease or diabetes significantly increase the risk of death from COVID-19. Symptoms such as dyspnea and fever were also associated with fatal outcomes. In addition, the study highlighted a higher mortality rate among men and underscored the severe impact of COVID-19 on individuals with preexisting health conditions (OLIVEIRA et al., 2021).

(SNIDER; PATEL; MCBEAN, 2021) aims to identify the main risk factors that influence mortality from COVID-19 in patients in Ontario, Canada, using XGBoost. The goal is to provide insights that can help optimize healthcare resources and formulate more effective public policies. Data from 57,390 patients diagnosed with COVID-19 were collected through the Ontario Health Data Platform. This data included demographic, epidemiological, and comorbidity information. After collection, the data were divided into a training set and a test set, with the first used to train the model and the second to validate its effectiveness. XGBoost was used to analyze the data and identify patterns that could predict mortality risk. SHAP Values were used to interpret the model results and understand how each variable impacted mortality prediction.

The study indicated that patient age was the most critical factor, with significantly higher risks for older groups. The test date was also relevant, suggesting that individuals diagnosed in more advanced stages of the pandemic had lower mortality risks, possibly due to advances in treatment and management of the disease. Other variables such as sex, income, and ethnicity were also highlighted as influential, with lower-income and more ethnically diverse groups showing greater vulnerability (SNIDER; PATEL; MCBEAN, 2021).

(WEN et al., 2022) aims to investigate how non-pharmaceutical interventions (NPIs) influence COVID-19 mortality before and after the implementation of mass vaccination campaigns. The study analyzed data from 34 countries, focusing on the relationship between the stringency of NPIs and mortality rates at different stages of the pandemic. Initially, the study collected and analyzed data on implementing NPIs and COVID-19 mortality rates. Using the Random Forest algorithm, the study modeled the complexity of the relationships between the variables. To interpret the results of the Random Forest model and assess the importance of each NPI in reducing mortality, the study employed SHAP Values. These values provide a detailed and quantitative explanation of the impact of each intervention, allowing a deeper understanding of which measures are most effective in different contexts.

The results indicated that, early in the pandemic, stricter NPIs were strongly associated with a significant reduction in mortality. However, after vaccination was implemented, the influence of these interventions on mortality decreased. SHAP analyses highlighted that measures such as travel restrictions and cancellation of public events were initially crucial, but their importance decreased as vaccination progressed. The study concludes that, although vaccination has reduced the need for strict NPIs, these interventions still play a vital role, especially in areas with incomplete vaccination coverage. The research highlights the need to adapt public health policies as the epidemiological situation evolves, using a combination of vaccination and NPIs to manage the pandemic effectively (WEN et al., 2022).

(ZHOU et al., 2023) main objective is to investigate how booster vaccination influences the age-adjusted case fatality rate of COVID-19 in 32 countries. To this end, the study used XGBoost and SHAP Values analysis to provide a more detailed interpretation of the factors involved. The study steps included collecting and analyzing global data on age-adjusted COVID-19 infections and deaths from February 2020 to February 2022. In addition, data on vaccination and virus variants were collected. The study calculated crude and age-adjusted case fatality rates and analyzed several explanatory variables across six dimensions: demographics, disease burden, and health services. XGBoost was used to model the age-adjusted case fatality rates. Feature selection was done through a recursive elimination process, and hyperparameter tuning was done through a grid search with cross-validation. The study used SHAP Values to interpret the models generated by XGBoost.

The results indicated significant variation in case fatality rates between countries and over time, especially with the introduction of new virus variants. Booster vaccination was found to play a crucial role in reducing case fatality, with positive impacts observed in simulations that increased booster vaccination coverage by 1 to 30%. The study also highlighted the importance of evidence-based public health strategies tailored to the specific characteristics of each country to combat the pandemic effectively (ZHOU et al., 2023).

3
Methodology

This study reports a retrospective analysis of COVID-19 mortality data in LAC countries. It was considered its evolution over time and examined how vaccination coverage, containment measures, economic support, and population mobility contributed to the evolution of deaths from COVID-19 in LAC countries, considering their socioeconomic differences. The construction process and the relevant steps of the method proposed by this work are illustrated in 3.1. The flowchart details the steps from obtaining and preparing the data to developing the final models, which estimate the contribution of the variables and evaluate the results with the experts. The only ones above each step are the numbers of the chapters of this work where each of them was detailed.

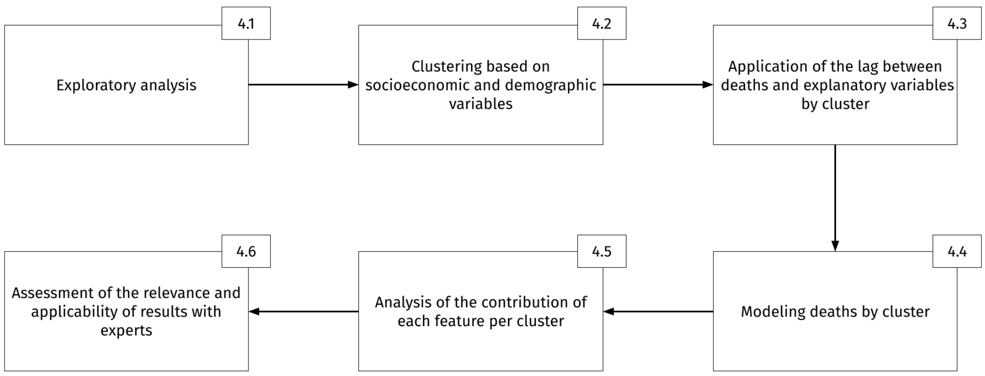

Figure 3.1: **Flowchart of research steps**

The data sources used in the study are described in the table 3.1.

Table 3.1: Description of the data sources used in this study

Data	Source	Source address
COVID-19 cases and deaths	Our World in Data	https://ourworldindata.org/
Containment measures during the COVID-19 pandemic	Oxford Covid-19 Government Response Tracker project	https://www.bsg.ox.ac.uk/research/covid-19-government-response-tracker
Population mobility trends during the COVID-19 pandemic	Google Mobility Data	www.google.com/covid19/mobility
Sociodemographic and economic data	Our World in Data, Human Development Reports, and The World Bank	Our World in Data: https://ourworldindata.org/ Human Development Reports: https://hdr.undp.org/data-center The World Bank: https://data.worldbank.org/indicator/

The study included data from January 1, 2020, to October 15, 2022. It was analyzed data from all 20 LAC countries (Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica, Cuba, Ecuador, El Salvador, Guatemala, Haiti, Honduras, Mexico, Nicaragua, Panama, Paraguay, Peru, Dominican Republic, Uruguay, and Venezuela).

The outcome is the COVID-19 mortality rate, calculated as the 7-day daily moving average of COVID-19 deaths per 100,000 population.

All analysis and modeling were performed in Python. Except for the mobility variables, all variables were standardized by the Z-Score method and then by the Min-Max standardization to keep the variable values between 0 and 1. The mobility variables were only standardized by the Z-Score method since these variables can assume negative values.

3.1 Clustering

According to (TAVARES; BETTI, 2021), people living in a country face multiple difficulties simultaneously. Therefore, when characterizing a country, it is important not to consider a single dimension (for example, high-income or low-income). Thus, the proposed method addresses various socioeconomic and demographic characteristics of countries through the clustering technique to reduce dimensionality of the data and create a single to classification to be used as an adjustment covariate in the modelling stage. Before clustering, an analysis was conducted to detect multicollinearity in the candidate variables for clustering using VIF (Variance Inflation Factor), where variables with a VIF greater than ten were removed (O'BRIEN, 2007). The K-Means(IKOTUN et al., 2023) and DBSCAN (ESTER et al., 1996) clustering methods were tested, and the silhouette index, the elbow method, and the Calinski-Harabasz score were tested to choose the number of clusters.

K-Means is a clustering algorithm used in machine learning and data analysis to partition a set of n observations into k clusters, where each observation belongs to the cluster with the closest centroid. K-Means aims to minimize the sum of the squared distances between the observations and the cluster centroids, resulting in a partition that maximizes the internal homogeneity of the clusters and the heterogeneity between them. The algorithm follows an iterative process that starts with initializing the centroids, assigning each observation to the closest cluster, and updating the centroids based on the means of the assigned observations. This process is repeated until convergence is achieved and the cluster assignments no longer change significantly (IKOTUN et al., 2023).

DBSCAN (Density-Based Spatial Clustering of Applications with Noise) is a density-based clustering algorithm used to identify arbitrary-shaped clusters in noisy datasets. Unlike methods such as K-Means, which require the number of clusters to be specified in advance, DBSCAN automatically identifies the number of clusters based on the local density of the data points (ESTER et al., 1996).

3.2 Day's lag and Poisson regression

To assess the effectiveness of interventions during the pandemic, such as the closure of schools and the percentage of the population that is fully vaccinated, it is important to lag deaths from interventions, as this lag represents the time between the implementation of the measures and the observation of the results of these measures on the number of deaths. This day's lag was calculated to find the ideal lag between the mortality rate and the interventions. Lags of 21 to 49 days (COSTA et al., 2023) were tested for each cluster. To define the best lag, Poisson regressions (COXE; WEST; AIKEN, 2009) were applied to each cluster, and each lag was tested. The best lag was chosen based on the AIC. After the lag, a multicollinearity analysis was carried out with the VIF in the intervention variables, and, again, variables with a VIF greater than ten were removed (O'BRIEN, 2007).

Poisson regression is a statistical technique used to model count data, where the dependent variable is a count of events that occur in a fixed time or space interval. Based on the Poisson distribution, this regression assumes that the mean of the count is equal to its variance, which makes it particularly suitable for count data that follow this property (COXE; WEST; AIKEN, 2009).

3.3 XGBoost

An XGBoost (Extreme Gradient Boosting) model (CHEN; GUESTRIN, 2016) was then applied, with the lagged COVID-19 mortality rate as its response variable and a 7-day moving average adjusted by the predominance of each COVID-19 variant per country. The intention was to remove the confounding effects of variants in analyzing the measures that most influenced mortality. The present work extensively experimented by varying its parameters to improve the proposed model's performance.

XGBoost belongs to Ensemble Learning models, which are based on training several simple models to produce a more robust final model. The algorithm builds decision trees in each iteration, where the models are no longer trained independently but sequentially, adjusting themselves based on previously trained models. XGBoost stands out for its regularization capacity, which helps to avoid overfitting, and for its computational efficiency, allowing parallel execution and handling sparse data (CHEN; GUESTRIN, 2016).

A Bayesian optimizer was used with cross-validation to optimize the hyperparameters of the XGBoost model. This method constructs a probabilistic model of the loss function concerning the hyperparameters and uses this model to make an informed selection of the next points to be evaluated (MOčKUS, 1975). An k-cross validation is incorporated into the process to ensure that the evaluation of the hyperparameters is robust and generalizable, minimizing the risk of overfitting. This iterative and adaptive process is particularly useful for XGBoost, which has a complex and high-dimensional hyperparameter space, including parameters such as learning rate, tree depth, regularization, and number of estimators.

3.4 SHAP Values

Finally, the SHAP Values technique (LUNDBERG; LEE, 2017) was applied to the results of the XGBoost modeling to understand how each feature of the model is associated with mortality from COVID-19.

SHAP Values (Shapley Additive Explanations) is a game-theoretic methodology for explaining the predictions of machine learning models, including XGBoost. Based on the concept of Shapley values, which assign each feature its marginal contribution to the model's output, SHAP Values provide a consistent and locally accurate decomposition of predictions. This is particularly useful for XGBoost, a model with low explainability due to its complexity and large number of hyperparameters. By applying SHAP Values, it is possible to interpret transparently and quantitatively the influence of each input variable on the model's output (LUNDBERG; LEE, 2017).

3.5 Assessment of the relevance and applicability of results with experts

In the research, experts evaluated the relevance and applicability of the results, aiming to ensure that the conclusions obtained were not only theoretically robust but also applicable and useful in a practical context. This validation with experts is crucial, as it allows the results to be interpreted based not only on theory but also on experience and practical knowledge, identifying possible limitations and opportunities for improvement that may not be evident through quantitative analysis alone. In addition, validation with experts ensures that the recommendations derived from the research are feasible and aligned with the needs and realities of the field of study (SHCHERBAKOV et al., 2014). The research was presented to 5 health field experts, who discussed the results and evaluated their relevance and employability. Individual sessions were held with each expert, each session lasting approximately 45 minutes. During the session, slides were presented describing the study's objective and goals and summarizing the results obtained. At the end, each expert answered a set of 4 questions for each period analyzed (pre- and post-vaccination periods). The questions were:

1. Were the results presented surprising/unexpected?
2. What factors do you believe led to these results?
3. To what extent do these results contribute to the formulation of public policies in Latin America?
4. If you were a manager, what guidelines would you create?

The objective was to create both questions that refer to the Likert scale, for quantifiable analysis, and open-ended questions, for qualitative analysis.

4
Results

4.1
Exploratory analysis

As seen in Fig 4.1, Peru had the most deaths per 100,000 pop. until October 15, 2022, followed by Brazil and Chile. The number of deaths in Peru differs significantly from other countries. The countries with the fewest deaths are Nicaragua, Haiti, and Venezuela, in that order.

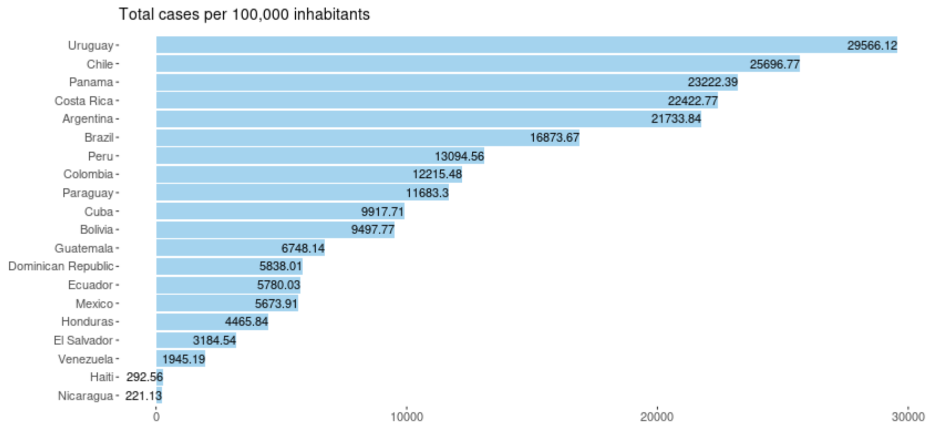

Figure 4.1: **Total deaths per 100,000 pop. by country until October 15, 2022.**

In Fig 4.2, it is possible to observe that each country had different waves of daily deaths and dimensions for this number of deaths.

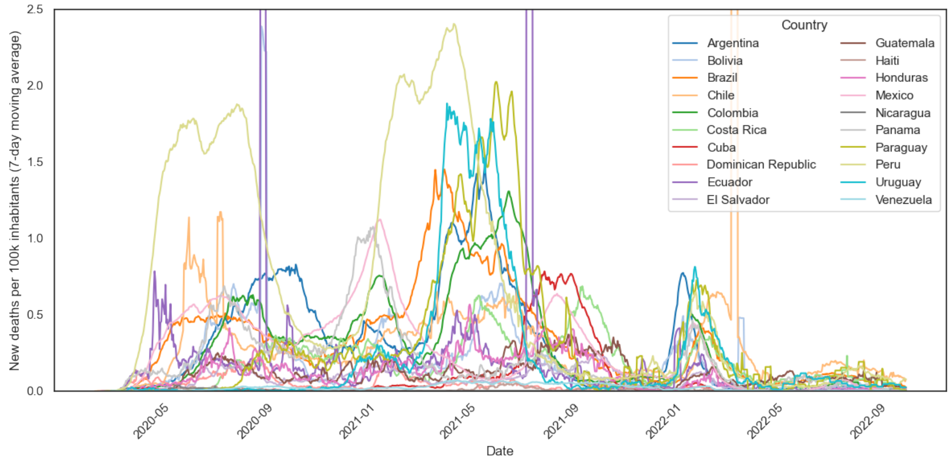

Figure 4.2: **New deaths per 100k pop. (7-day moving average) time series per LAC country.** Chile, Ecuador, and El Salvador have had unexpected jumps in deaths from COVID-19, and on specific days in the Our World in Data database.

The country with the largest share of the population living in extreme poverty and the largest share of the population in informal work is Haiti, as shown in Fig 4.3. The country with the greatest inequality according to the Gini index is Brazil. Meanwhile, the country with the lowest GDP per capita is Haiti. Among the ten socioeconomic and demographic variables, only Cuba and Venezuela have missing values or values equal to zero for some variables. The variables are:

- Cuba: share of informal employment, share of the population living in extreme poverty, Gini index, GDP per capita
- Venezuela: share of informal employment, share of the population living in extreme poverty

Furthermore, Cuba does not have Google data on population mobility, which is crucial for the research objective. Considering these points, Cuba and Venezuela were removed from the analysis.

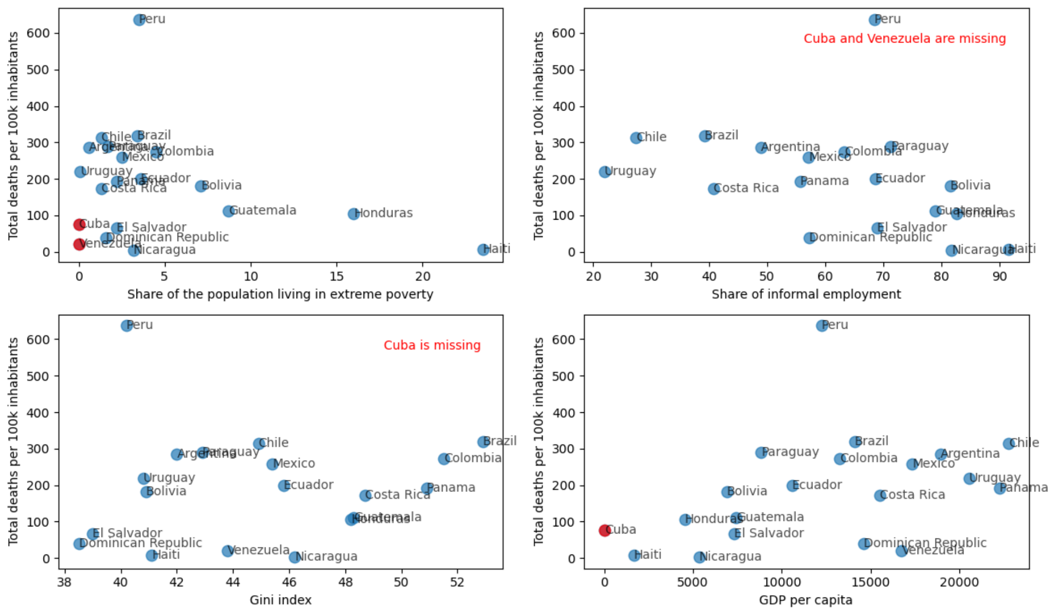

Figure 4.3: **Relationship among total deaths per 100,000 pop. per country and the share of the population living in extreme poverty, the share of informal employment, the Gini index, and GDP per capta.**

4.2
Clustering based on socioeconomic and demographic variables

Clustering was carried out to consider the different socioeconomic and demographic characteristics. However, before clustering, variables with VIF greater than ten were removed. Therefore, the variables selected for clustering were the Gini Index, average household size (number of members), median age of the population, share of the population living in extreme poverty, percentage of the population aged between 20 and 79 years who have type 1 or type 2 diabetes, population density. The variables removed were HDI, the average number of years of education received by people aged 25 and over, GDP per capita, and share of informal employment.

Based on the clustering methods tested, the method chosen was K-Means with 2 clusters as shown in Fig 4.4, where cluster 0 comprises Argentina, Brazil, Chile, Colombia, Costa Rica, Mexico, Panama, and Uruguay. In contrast, cluster 1 comprises Bolivia, Dominican Republic, Ecuador, El Salvador, Guatemala, Haiti, Honduras, Nicaragua, Paraguay, and Peru.

Figure 4.4: **Map of LAC countries by cluster. Cluster 0 is a less vulnerable cluster compared to Cluster 1.**

In Fig 4.5, it is possible to observe the distribution of clustering variables by cluster. The Mann-Whitney test indicated that, for a p-value of 0.05, the Gini index variables, the share of the population with diabetes, and population density do not have different medians. Furthermore, it is possible to observe that cluster 1 has a younger population than cluster 0, in addition to being the cluster with the largest population in extreme poverty and the largest average household size. In this way, cluster 0 is less vulnerable compared to cluster 1. Therefore, from now on, the present work will call cluster 0 less vulnerable countries and cluster 1 will be called more vulnerable countries.

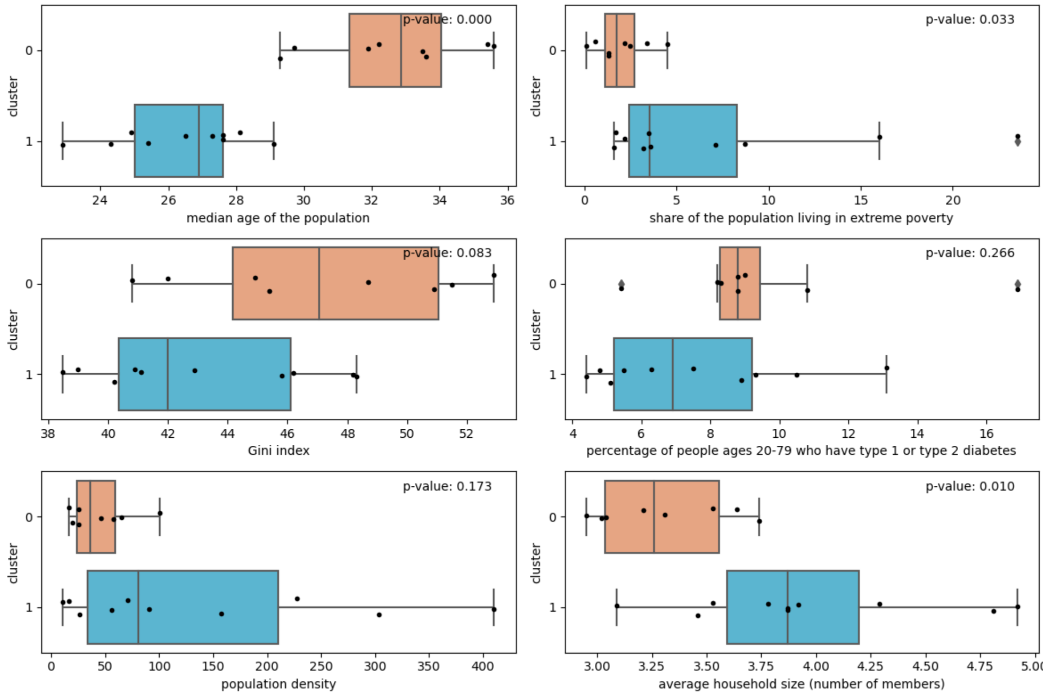

Figure 4.5: **Boxplot of socioeconomic and demographic characteristics by cluster. Cluster 0 is a less vulnerable cluster compared to Cluster 1.**

4.3 Application of the lag between deaths and explanatory variables by cluster

It is important to lag deaths from interventions to assess the effectiveness of interventions during the pandemic (LI et al., 2022; COSTA; ROHLEDER; BOZORGMEHR, 2023). Therefore, a lag was applied between deaths and explanatory variables, which are interventions per cluster, where 21 to 49 days lags were tested for each cluster. The lags chosen for each cluster were:

– Less vulnerable cluster: 49 days

– More vulnerable cluster: 21 days

A possible interpretation for these lags is that more vulnerable cluster interventions had faster results than less vulnerable cluster interventions.

4.4 Modeling deaths by cluster

The next step was to select the variables that would be used to explain deaths from COVID-19. For this, again, variables with VIF greater than ten were removed. In the end, the variables selected for modeling were school closures, restrictions on the size of meetings between people, closure of public transport, requirements to stay at home, restrictions on internal movement, debt/contract relief for families, income support for people who have lost their job or who cannot work, variation in mobility in supermarkets and pharmacies, variation in mobility in parks, variation in mobility in public transport, variation in mobility in places of work, variation in mobility in residential locations, percentage of the population vaccinated with at least two doses. Furthermore, the percentage of predominance of each variant per country was also included in the model to remove confounding effects associated with COVID-19 variants from the analysis.

In total, was obtained four models and the SHAP Values technique was applied to the four models obtained:

1. Pre-vaccination period for the less vulnerable countries
2. Pre-vaccination period for the less vulnerable countries
3. Post-vaccination period for the more vulnerable countries
4. Post-vaccination period for the more vulnerable countries

4.5 Analysis of the contribution of each feature per cluster

4.5.1 Pre-vaccination period

SHAP Values show that, when comparing less vulnerable countries with more vulnerable countries in the pre-vaccination period, it is possible to see from Fig 4.6 that the most important variables in less vulnerable countries were containment measures, while in more vulnerable countries, they were mobility variables. Furthermore, economic support variables were more important in less vulnerable than in more vulnerable countries.

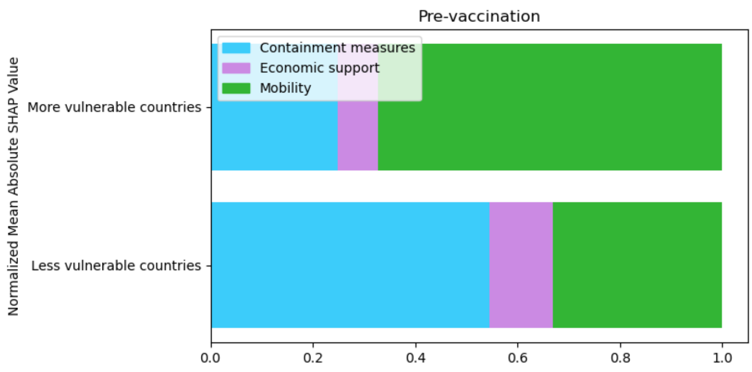

Figure 4.6: **Normalized mean absolute SHAP Value by cluster for the pre-vaccination period.**

Although containment measures were the most influential variables in less vulnerable countries, their relevance was mainly driven by the closure of public transport and schools, `C5_Close public transport` and `C1_School closing_Agrup` respectively, as can be seen in Fig 4.7. The most important variables for the most vulnerable countries were the variation in mobility in residential areas and the closure of public transport, `residential_percent_change_from_baseline` and `C5_Close public transport` respectively. The closure of public transport was relevant in both clusters.

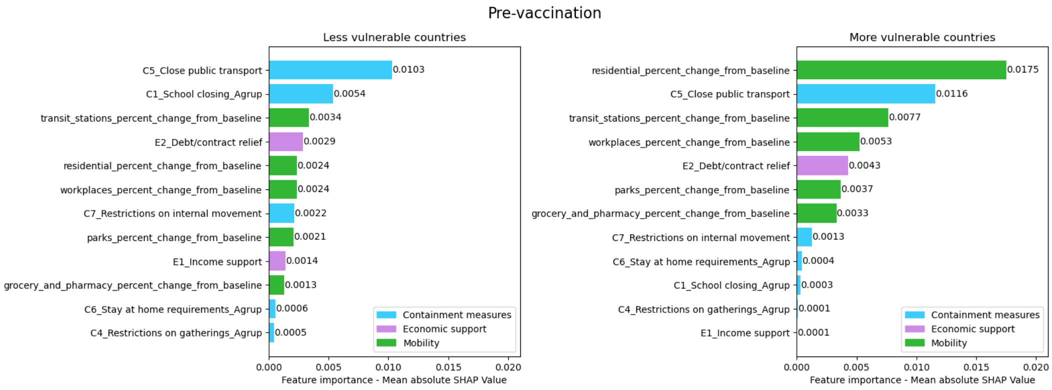

Figure 4.7: **Mean absolute SHAP Value by cluster for pre-vaccination period**

Fig 4.8 shows that school closures, `C1_School closing_Agrup`, are the second most important variable in less vulnerable countries, but it drops eight positions in the importance ranking of more vulnerable countries. The mobility variation variable in transport stations, `transit_stations_percent_change_from_baseline`, has the same position in the importance ranking of both clusters. Both economic support variables are less important in more vulnerable countries than in less vulnerable ones.

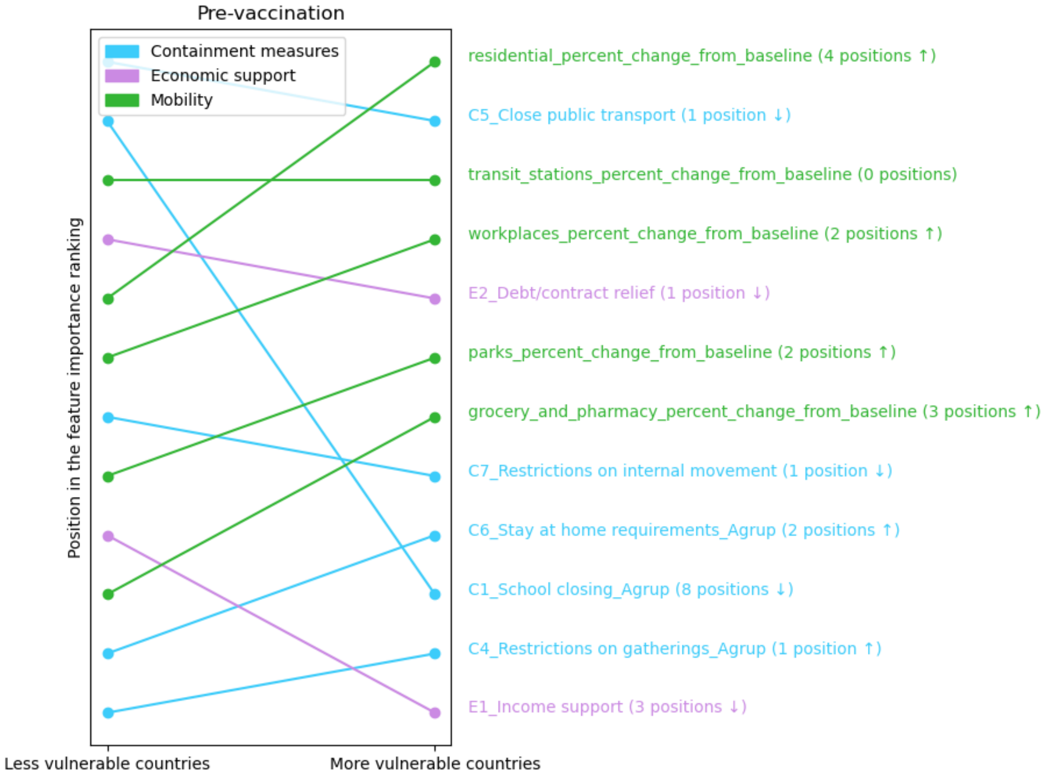

Figure 4.8: **Position in the feature importance ranking by cluster for pre-vaccination period**

In Fig 4.9, it is possible to observe that in the pre-vaccination period, debt relief and economic support, `E2_Debt/contract relief` and `E1_Income support` respectively, are associated with reducing mortality from COVID-19 in less vulnerable countries. Some counterintuitive results indicate that the closure of public transport, the closure of schools, the increase in mobility in public transport stations, the increase in mobility in work areas, parks, stores, and pharmacies, and the restrictions on the size of meetings, `C5_Close public transport`, `C1_School closing_Agrup`, `transit_stations_percent_change_from_baseline`, `workplaces_percent_change_from_baseline`, `parks_percent_change_from_baseline`, `grocery_and_pharmacy_percent_change_from_baseline` and `C4_Restrictions on gatherings_Agrup` respectively, between people are associated with increased mortality.

For more vulnerable countries, unlike less vulnerable countries, the debt relief variable, `E2_Debt/contract relief`, is associated with increased mortality. The variation in mobility in residential areas, `residential_percent_change_from_baseline`, is the most important feature of the model, but it is always associated with an increase in mortality. The variation in mobility at transport stations, `transit_stations_percent_change_from_baseline`, is mostly associated with a reduction in mortality, as is the variation in mobility in parks, markets, and pharmacies, `parks_percent_change_from_baseline` and `grocery_and_pharmacy_percent_change_from_baseline` respectively.

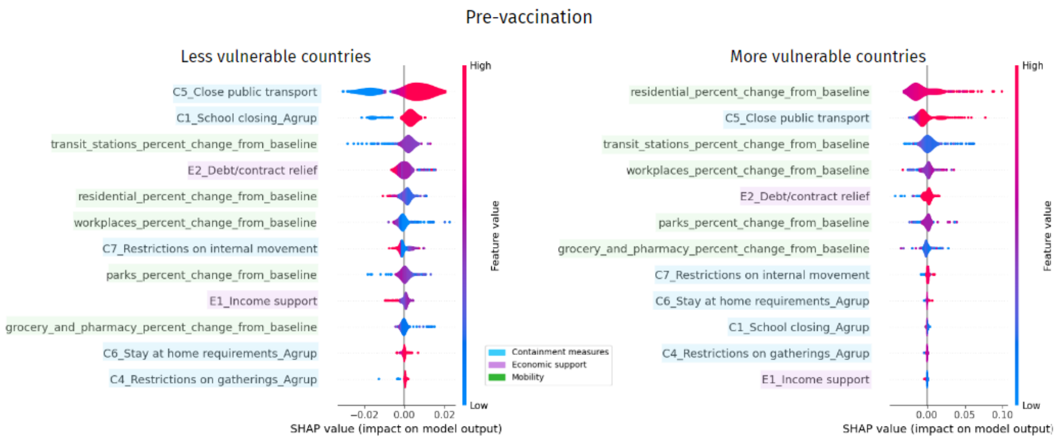

Figure 4.9: **SHAP values by cluster for pre-vaccination period** On the y-axis are the features, and on the x-axis are the SHAP values. Each point represents a data instance, and the color represents the feature value. The plot shows only the nine most relevant features according to feature importance.

4.5.2 Post-vaccination period

Fig 4.10 shows that the vaccination coverage variable in the post-vaccination period was the most important in less vulnerable countries. However, the most important variables for more vulnerable countries were the containment measure variables, followed by the mobility variables.

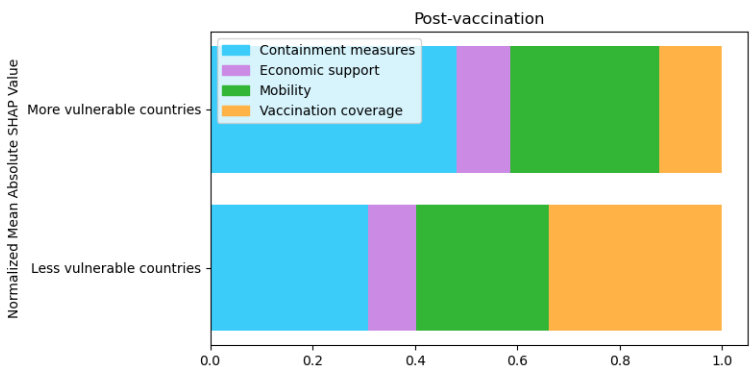

Figure 4.10: **Normalized mean absolute SHAP Value by cluster for the post-vaccination period.**

As seen in Fig 4.11, the vaccination variable, `people_at_least_two_doses_vaccinated_per_p` in the less vulnerable countries is much more relevant than the others. Although in more vulnerable countries, the type of variable with the most importance is containment measures, the variable that drives this importance is the requirement to stay at home, `C6_Stay at home requirements_Agrup`. In more vulnerable countries, the variables of stay-at-home requirements are more important than the vaccination coverage variable.

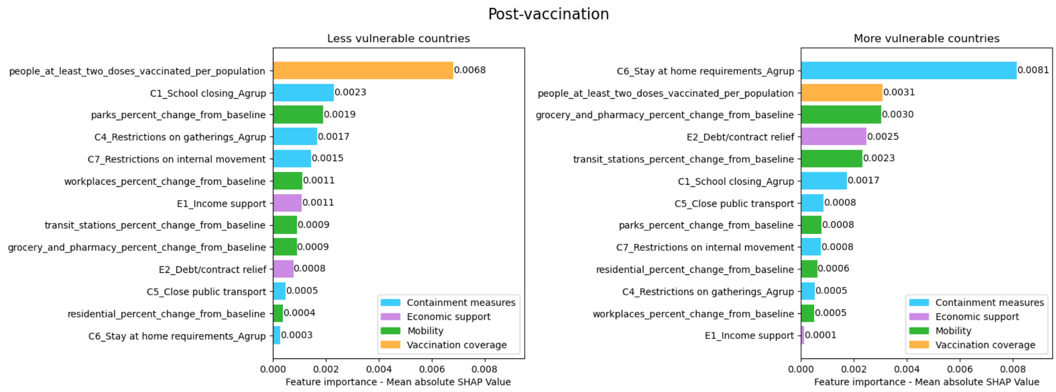

Figure 4.11: **Mean absolute SHAP Value by cluster for post-vaccination period**

The economic support variable, `E2_Debt/contract relief`, drops six positions in the importance ranking of more vulnerable countries. The stay-at-home requirement variable, `C6_Stay at home requirements_Agrup`, is the most important for the most vulnerable countries and the least important for the least vulnerable countries. Furthermore, the vaccination variable, `people_at_least_two_doses_vaccinated_per_population`, drops one position in the importance ranking of more vulnerable countries, as seen in Fig 4.12.

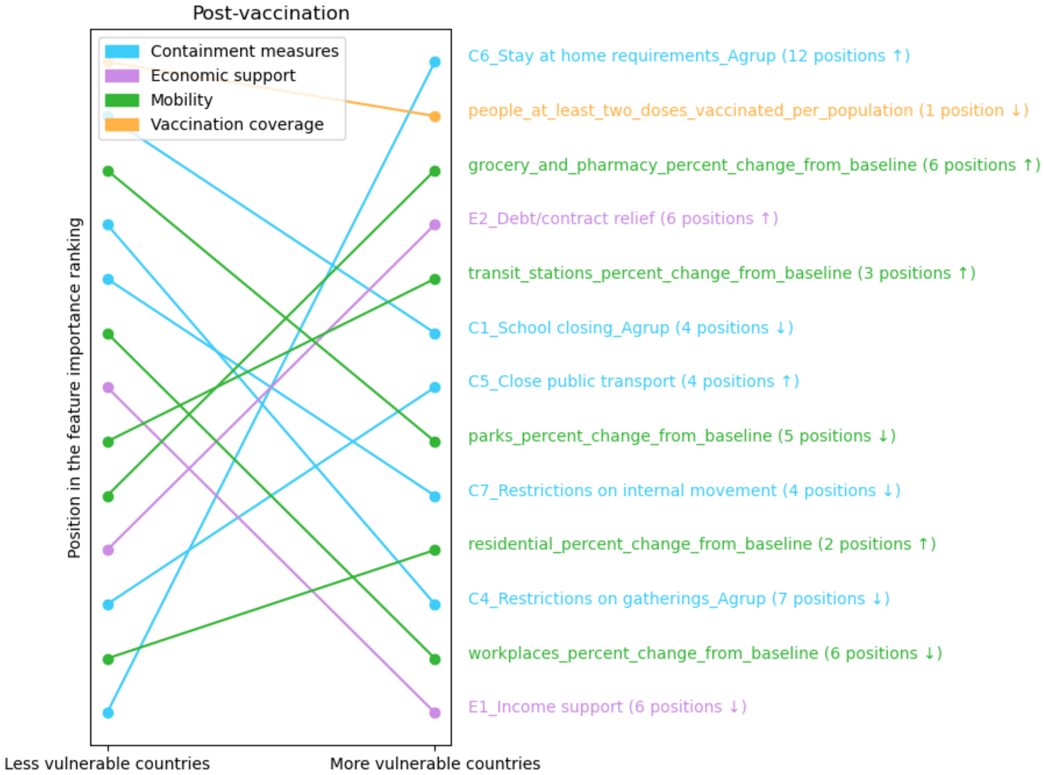

Figure 4.12: **Position in the feature importance ranking by cluster for post-vaccination period**

In Fig 4.13, it is possible to observe that in the post-vaccination period, the percentage of the population vaccinated with at least two doses, `people_at_least_two_doses_vaccinated_per_population`, is associated with a reduction in mortality from COVID-19 in both less and more vulnerable countries. However, for more vulnerable countries, it is possible to observe that the requirements to stay at home, `C6_Stay at home requirements_Agrup`, are more relevant than vaccination and are associated with increased mortality.

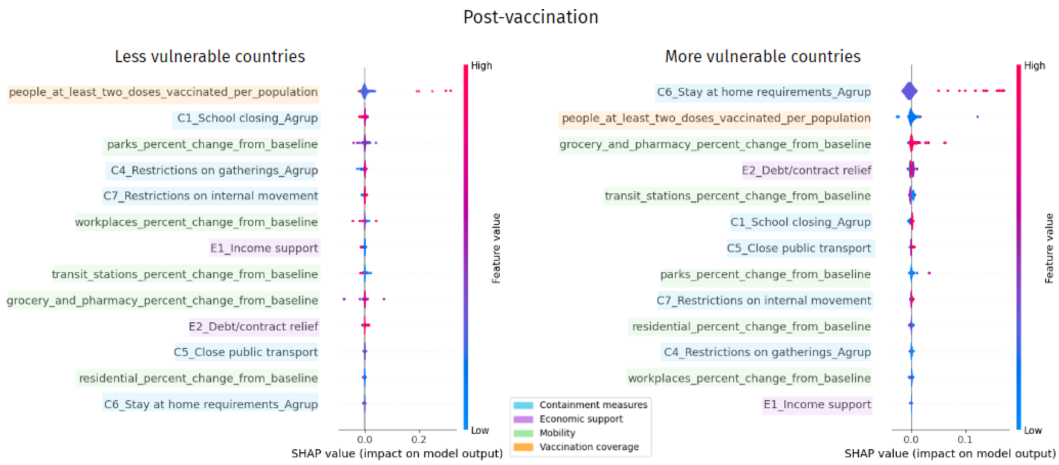

Figure 4.13: **SHAP values by cluster for post-vaccination period** On the y-axis are the features, and on the x-axis are the SHAP values. Each point represents a data instance, and the color represents the feature value. The plot shows only the nine most relevant features according to feature importance.

Applying the SHAP Values technique to the four models revealed significant insights into the factors influencing COVID-19 mortality in different cluster contexts. It was possible to observe a marked distinction between less and more vulnerable countries in the pre-vaccination period. While containment measures emerged as the most important variables in less vulnerable countries, more vulnerable countries were mainly characterized by mobility variables. Although containment measures proved crucial in less vulnerable countries, their importance was largely driven by the closure of public transport and schools.

However, when analyzing the post-vaccination period, it was possible to observe a significant change in dynamics. Vaccination coverage emerged as the most important variable in less vulnerable countries, reflecting the tangible benefits of vaccination in reducing mortality. Conversely, more vulnerable countries continued to be influenced mainly by containment measures and mobility variables. Some hypotheses for this are that less vulnerable countries have greater access to vaccines or that they have greater vaccination coverage than more vulnerable countries.

Some findings were unexpected; for example, during the pre-vaccination phase in both less and more vulnerable countries, there was a correlation between increased deaths and public transport closures. A limitation of the current model is that it does not account for causality between variables. However, the lag was used to model the potential time course of the intervention effect.

Despite some counterintuitive results, the model presented two that made sense given the history of COVID-19. It is known that the greater the percentage of the population with at least two doses of vaccine, the lower the number of deaths. There is also a hypothesis that when the government helps the population financially, people tend to stay home more, reducing the number of deaths.

4.6 Assessment of the relevance and applicability of results with experts

Public health experts were invited to participate in panel discussions to assess the results of this research. The experts of these panels revealed a number of valuable insights into the complexity and challenges faced in public health policy-making. One of the most discussed points was the relationship between correlation and causation. Experts highlighted the difficulty of establishing robust causal relationships from correlations observed in aggregated data. Three experts emphasized that the accuracy and validity of measures used to capture phenomena such as mobility and adherence to containment policies are crucial to avoid biases that could distort the results. (FABIO et al., 2021) addresses the complexity and diversity of innovative lockdown measures during the COVID-19 pandemic. The authors highlight that while several localities have lockdown measures in place, these measures vary significantly in terms of definition and application, making it difficult to compare and assess their effectiveness across different regions and countries.

Another recurring theme in the discussions was the heterogeneity among regions and how this influences the effectiveness of public health policies. Four experts agreed that socioeconomic and demographic diversity across regions makes comparing policies and measures a significant challenge. One of the experts highlighted that "if there are already huge differences within Brazil, imagine comparing several countries", highlighting the complexity of generalizing the results of public health policies. In addition, it was mentioned that adherence to containment measures varied significantly among regions, which may explain the unexpected results observed in some studies, as shown in (FABIO et al., 2021). The need to consider specific regional characteristics when formulating public health policies was widely recognized.

Vaccination was another central topic in the discussions. All experts agreed that vaccination played a crucial role in reducing mortality from COVID-19, especially in less vulnerable countries. However, it was also highlighted that the availability of and the access to vaccines varied significantly among regions, influencing the effectiveness of vaccination campaigns. Reducing inequity in access to the vaccine was highlighted as a priority for future vaccination campaigns by 3 experts.

Economic support during the pandemic was also widely discussed. Two experts recognized that economic support measures, such as debt relief and income distribution, were essential to mitigate the economic impacts of the pandemic, especially in less vulnerable countries. However, they observed that the impact of these measures varied among regions. One expert mentioned that "the more economic support I provide, the lower my mortality in these less vulnerable countries is", highlighting the importance of adapting economic support policies to the specific needs of each region.

All experts indicated that the research results added value to the current literature, since they brought to light points that are not intuitive and that do not simply reproduce the results already known. Given the novelty of the results, the experts do not feel comfortable stating that they would adopt these results individually to define new public policies. However, they believe that it is worth deepening the research further in the direction of formulating public policies. One expert said, "I do not think it would be possible to use this result, for example, to formulate a policy. I think it brought a new perspective to the problem; all the points you raised here are points of conflict in the literature."

In summary, all experts pointed out that the most interesting results as a public health policy were that vaccination coverage was associated with a reduction in COVID-19 mortality in both groups of countries and that economic support was associated with a decrease in COVID-19 mortality only in the least vulnerable countries, which implies that perhaps these countries were the ones that had the financial means to offer greater support. The expert discussions revealed the complexity and challenges in formulating public health policies from the analysis results of the COVID-19 pandemic. The need for careful and contextualized analysis that considers regional heterogeneity and the relationship between correlation and causality was widely recognized. In addition, the importance of vaccination as an essential tool to mitigate the impacts of the pandemic was highlighted.

5
Conclusion and future work

The first objective of the research was to understand which LAC countries are similar and how they are similar according to their socioeconomic and demographic characteristics. It was possible to detect two distinct groups of countries, where one group may have characteristics of greater vulnerability than the other group.

For the second objective of the research, the best interval of days between interventions (vaccination coverage, containment measures, economic support and population mobility) and mortality from COVID-19 was determined. This interval was 49 days for the least vulnerable countries and 21 days for the most vulnerable.

For the third research objective, it was possible to identify which interventions are most important for understanding COVID-19 mortality and how they are associated with COVID-19 mortality in two distinct periods of the pandemic: pre-vaccination and post-vaccination. In the pre-vaccination period, containment measures were the most important interventions for mortality in the least vulnerable countries. In contrast, for the most vulnerable countries, they were features of variation in population mobility. In the post-vaccination period, the most important interventions for mortality in the least vulnerable countries were vaccination coverage, while the most vulnerable countries were more impacted by containment measures.

However, the additional objective of proposing public health guidelines for future respiratory diseases was not achieved. Despite this, the experts state that the results presented were unexpected and that they add new perspectives compared to the current literature, which can generate new discussions to enrich knowledge on the subject. The results already known and that are in agreement with what is in the literature were that vaccination coverage was associated with a reduction in COVID-19 mortality in both groups of countries and that economic support was associated with a decrease in COVID-19 mortality only in the least vulnerable countries, which implies that perhaps these countries were the ones that had the financial means to provide greater support. In view of the unexpected and counterintuitive results, as future work it is interesting to apply XAI (Explainable AI) and causal inference analysis techniques, in order to deepen the understanding of the results found.

Although the research identified significant associations between the variables, it is essential to recognize that correlation does not imply causality.

The observed relationships may be influenced by uncontrolled confounders or omitted variables that simultaneously affect the variables of interest. Causal inference analysis would help distinguish between associations and causality. Furthermore, it is difficult to isolate the effects of different interventions due to the simultaneity of the measures.

One point of discussion in the research is the heterogeneity of the regions studied and its impact on the generalizability of the results. Latin America is characterized by substantial socioeconomic, cultural and geographic diversity, which may influence the observed effects of the variables analyzed differently. This heterogeneity may limit the generalizability of the results obtained to the entire region, since conclusions derived from a specific context may not directly apply to other contexts with different characteristics. For example, public health policies that have proven effective in a country with robust infrastructure may not have the same impact in countries with underfunded health systems.

6
Bibliography

BARGAIN, O.; AMINJONOV, U. Poverty and covid-19 in africa and latin america. **World Development**, Pergamon, v. 142, p. 105422, 6 2021. ISSN 0305-750X.

BASTOS, L. S. et al. Primary healthcare protects vulnerable populations from inequity in covid-19 vaccination: An ecological analysis of nationwide data from brazil. **The Lancet Regional Health - Americas**, Elsevier, v. 14, p. 100335, 10 2022. ISSN 2667-193X.

BENNETT, M. All things equal? heterogeneity in policy effectiveness against covid-19 spread in chile. **World Development**, Pergamon, v. 137, p. 105208, 1 2021. ISSN 0305-750X.

BURKI, T. Covid-19 in latin america. **The Lancet Infectious Diseases**, Elsevier, v. 20, p. 547–548, 5 2020. ISSN 1473-3099.

CHAN, I. L. et al. Promoting immunization equity in latin america and the caribbean: Case studies, lessons learned, and their implication for covid-19 vaccine equity. **Vaccine**, Elsevier, v. 40, p. 1977–1986, 3 2022. ISSN 0264-410X.

CHEN, T.; GUESTRIN, C. Xgboost. In: . [S.l.]: ACM, 2016. p. 785–794. ISBN 9781450342322.

COSTA, D. et al. Impact of non-pharmaceutical interventions on covid-19 incidence and deaths: cross-national natural experiment in 32 european countries. **medRxiv**, Cold Spring Harbor Laboratory Press, p. 2022.07.11.22277491, 7 2023.

COSTA, D.; ROHLEDER, S.; BOZORGMEHR, m. K. Impact of non-pharmaceutical interventions on covid-19 incidence and deaths: cross-national natural experiment in 32 european countries. **medRxiv**, Cold Spring Harbor Laboratory Press, 2023.

COXE, S.; WEST, S. G.; AIKEN, L. S. The analysis of count data: A gentle introduction to poisson regression and its alternatives. **Journal of Personality Assessment**, v. 91, p. 121–136, 2 2009. ISSN 0022-3891.

ESTER, M. et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In: **Proceedings of the Second International Conference on Knowledge Discovery and Data Mining**. [S.l.]: AAAI Press, 1996. (KDD'96), p. 226–231.

FABIO, I. et al. "what exactly is a lockdown?": Towards an ontology-based modeling of lockdown interventions during the covid-19 pandemic. In: FARINELLI, F. et al. (Ed.). **ONTOBRAS 2021: Ontology Research in Brazil**. [S.l.]: CEUR, 2021. (CEUR Workshop Proceedings), p. 151–164.

FAUCI, A. S.; LANE, H. C.; REDFIELD, R. R. Covid-19-navigating the uncharted. **The New England Journal of Medicine**, 2020.

GARCIA, P. J. et al. Covid-19 response in latin america. **The American Journal of Tropical Medicine and Hygiene**, The American Society of Tropical Medicine and Hygiene, v. 103, p. 1765, 11 2020. ISSN 14761645.

IKOTUN, A. M. et al. K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. **Information Sciences**, v. 622, p. 178–210, 2023. ISSN 0020-0255.

KWOK, K. O. et al. Epidemic models of contact tracing: Systematic review of transmission studies of severe acute respiratory syndrome and middle east respiratory syndrome. **Computational and Structural Biotechnology Journal**, v. 17, p. 186–194, 2019. ISSN 20010370.

LI, H. et al. Effects of vaccination and non-pharmaceutical interventions and their lag times on the covid-19 pandemic: Comparison of eight countries. **PLOS Neglected Tropical Diseases**, v. 16, p. e0010101, 1 2022. ISSN 1935-2735.

LITEWKA, S. G.; HEITMAN, E. Latin american healthcare systems in times of pandemic. **Developing World Bioethics**, John Wiley Sons, Ltd, v. 20, p. 69–73, 6 2020. ISSN 1471-8847.

LUNDBERG, S. M.; LEE, S. A unified approach to interpreting model predictions. **CoRR**, abs/1705.07874, 2017.

MARTINEZ-VALLE, A. Public health matters: why is latin america struggling in addressing the pandemic? **Journal of Public Health Policy**, Nature Publishing Group, v. 42, p. 27, 3 2021. ISSN 1745655X.

MOčKUS, J. On bayesian methods for seeking the extremum. **Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)**, Springer, Berlin, Heidelberg, v. 27 LNCS, p. 400–404, 1975. ISSN 1611-3349.

OLIVEIRA, M. C. et al. Fatores associados ao óbito em casos confirmados de covid-19 no estado do rio de janeiro. **BMC Infectious Diseases**, BMC, v. 21, 12 2021. ISSN 14712334.

O'BRIEN, R. M. A caution regarding rules of thumb for variance inflation factors. **Quality & Quantity**, v. 41, p. 673–690, 9 2007. ISSN 0033-5177.

PABLOS-MéNDEZ, A. et al. Covid-19 in latin america. **BMJ**, British Medical Journal Publishing Group, v. 370, 7 2020. ISSN 1756-1833.

PANIZ-MONDOLFI, A. E. et al. The arrival of sars-cov-2 in venezuela. **The Lancet**, v. 395, p. e85–e86, 2020. ISSN 0140-6736.

PERES, I. T. et al. Sociodemographic factors associated with covid-19 in-hospital mortality in brazil. **Public Health**, W.B. Saunders, v. 192, p. 15–20, 3 2021. ISSN 0033-3506.

SHCHERBAKOV, M. et al. **Lean Data Science Research Life Cycle: A Concept for Data Analysis Software Development**. [S.l.]: Springer International Publishing, 2014. 708-716 p.

SNIDER, B.; PATEL, B.; MCBEAN, E. Insights into co-morbidity and other risk factors related to covid-19 within ontario, canada. **Frontiers in Artificial Intelligence**, Frontiers Media S.A., v. 4, p. 85, 9 2021. ISSN 26248212.

TAVARES, F. F.; BETTI, G. The pandemic of poverty, vulnerability, and covid-19: Evidence from a fuzzy multidimensional analysis of deprivations in brazil. **World Development**, Pergamon, v. 139, p. 105307, 3 2021. ISSN 0305-750X.

WEN, H. et al. Non-pharmacological interventions of travel restrictions and cancelation of public events had a major reductive mortality affect during pre-vaccination coronavirus disease 2019 period. **Frontiers in Medicine**, Frontiers Media S.A., v. 9, p. 2463, 8 2022. ISSN 2296858X.

ZHOU, C. et al. The role of booster vaccination in decreasing covid-19 age-adjusted case fatality rate: Evidence from 32 countries. **Frontiers in public health**, NLM (Medline), v. 11, p. 1150095, 4 2023. ISSN 22962565.