Analyzing and Simulating COVID-19 Spread: Insights for Policy Implications in the U.S.

Introduction

Covid-19, for the first time, was detected in Wuhan, Hubei province, China in 2019. The unstoppable and unexpected expansion of this infectious disease, on March 11, 2020 made the WHO declare a state of pandemic. It had a widespread impact on nearly every nation, affecting over 50 million individuals globally and was acknowledged as the catalyst for the most significant economic crisis since the conclusion of World War II. This crisis forced governments to navigate through an environment marked by profound uncertainty, as they grappled with complex decisions that balanced the demands of public health, economic stability, and societal well-being. Since there were no prior knowledge and drugs to stop and protect people from this particular virus, the governments have taken various policy responses and concentrated all the efforts on preventive methods and measures such as distancing, social isolation, strong limitation on travel and even closing borders (Davahli et al., 2020; Gursoy and Chi, 2020; OECD, 2020).

The Secretary of Health and Human Services (HHS) of the United States declared a public health emergency on January 31, 2020, under section 319 of the Public Health Service Act (42 U.S.C. 247d), in response to COVID-19. The record-long United States economic expansion came to an end as a result of the COVID-19 pandemic, with forecasts of a deep recession in 2020 (CEPAL, 2020).

This policy paper seeks to provide a comprehensive analysis of the spread of COVID-19 within the United States, examining the multifaceted factors contributing to its dissemination. Additionally, this paper introduces a mathematical contagion model, designed to enhance our understanding of the pandemic's trajectory and simulate policy interventions. Specifically, the model will allow us to assess the implications of introducing a vaccination policy at a specific period.

Purpose and Scope

This policy paper aims to fulfill several key objectives:

- **Analysis of COVID-19 Spread:** To comprehensively analyze the dynamics and correlations of COVID-19 spread in the United States, including regional variations and the impact of policy responses.
- **Analysis of test results:** To understand the impact of testing overall and eliminate any specific state cause that can also have the same effect as testing.
- **Contagion Model and Policy Simulation:** To introduce and implement a mathematical contagion model designed to enhance our understanding of COVID-19 spread, inform policy decisions, and forecast potential future scenarios. Also, to simulate the effects of policy interventions using the contagion model.
- **Recommendations:** To propose evidence-based policy recommendations for managing and mitigating the ongoing COVID-19 crisis, informed by the insights gained from the contagion model.

In pursuit of these objectives, this policy paper will draw upon a comprehensive review of available data, scientific research, policy documents, and expert insights. It is our hope that the findings and recommendations presented herein will contribute to informed decision-making and the development of effective strategies to combat the COVID-19 pandemic in the United States.

Data

This policy paper leverages two critical datasets to inform its analysis and modeling of COVID-19 spread in the United States.

1. COVID-19 State Data Set: This dataset encompasses a wide range of variables related to the COVID-19 pandemic within the United States. It includes state-specific information such as the number of tested individuals, infected cases, deaths, population demographics, healthcare infrastructure, economic indicators, and more. These data points are essential for assessing the pandemic's impact on various aspects of society, understanding the disparities in disease spread, and guiding policy recommendations.

2. COVID-19 Daily Data Set: This dataset provides daily updates on COVID-19 statistics, offering a granular view of the pandemic's progression over time. It includes information on daily case counts, hospitalizations, testing rates, mortality rates, and other relevant metrics. The temporal nature of this dataset enables us to analyze the dynamic spread of the virus and evaluate the effectiveness of policy interventions as they unfolded.

Analysis of COVID-19 Spread

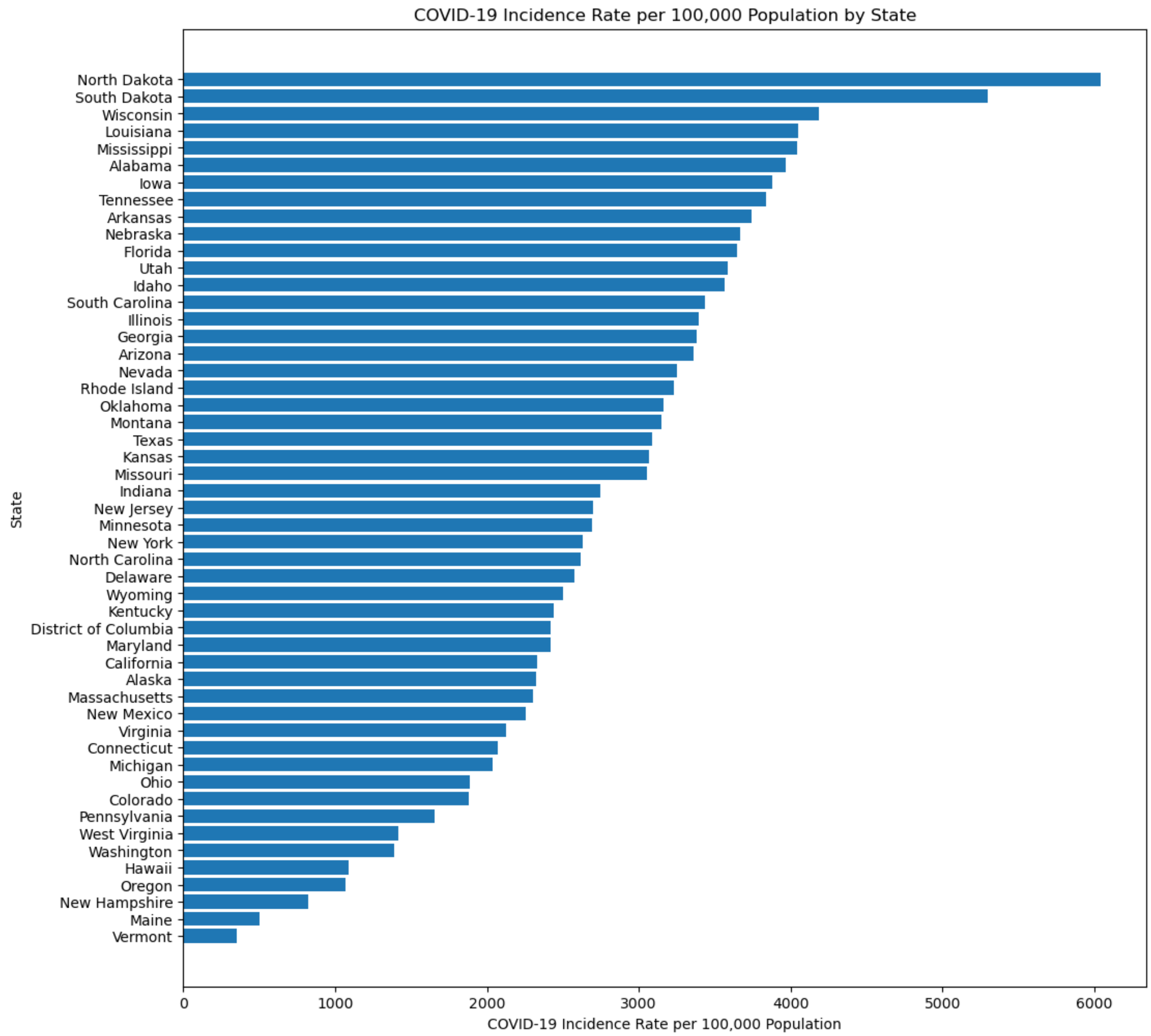

Figure 1. COVID-19 Incidences per 100k Population by State

In Figure 1, we present a visual representation of the COVID-19 incidence rate per 100,000 people across various states in the United States. This map illustrates how the pandemic has affected different regions, highlighting the varying degrees of infection rates.

This figure encapsulates the geographical disparities in COVID-19 incidence, offering a snapshot of the pandemic's impact on different states. It sets the stage for a detailed examination of the factors contributing to these variations and their implications for public health strategies and interventions.

Evolution of COVID-19 Cases Over Time

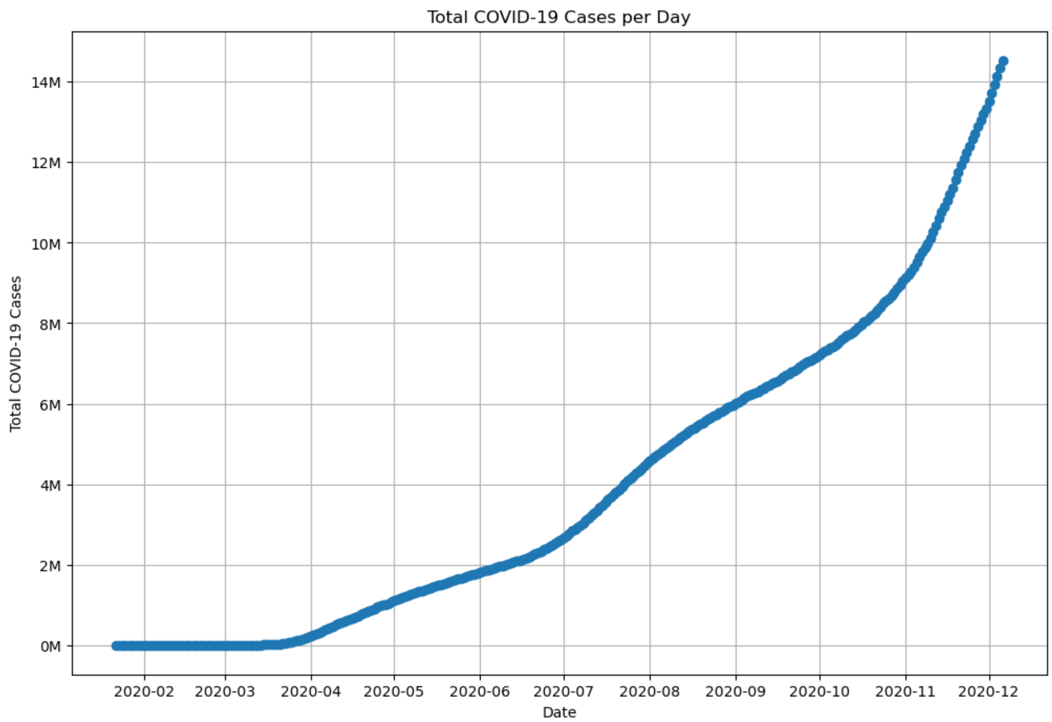

Figure 2. Total COVID-19 Cases per Day

Figure 2 illustrates the increase in the number of positive COVID-19 cases over time. This graph clearly demonstrates that the pandemic has expanded over the course of its progression. The rising trend is indicative of the virus's ability to spread and infect a growing portion of the population. By the end of December 2020 after a steady growth during the year, the peak value of cases per day reached around 14 millions of cases per day. The slope in the growth of these cases can be the representations of the different covid outbreaks, that coincide to where the line is growing in a steeper fashion. However, this graph, on its own, may not provide a complete picture of the pandemic's dynamics.

Figure 3 reveals a crucial aspect of the pandemic response—increasing testing efforts over time. As the pandemic unfolded, there was a notable escalation in the number of tests conducted. This can be attributed to several factors, including improved testing infrastructure, increased awareness of the importance of testing, and the need to identify and isolate infected individuals promptly. The rise in testing is not only a reflection of increased awareness but also a proactive strategy to detect cases earlier. The number of tests taken, by the end of December 2020 reached a peak value that exceeded 200 millions tests per day.

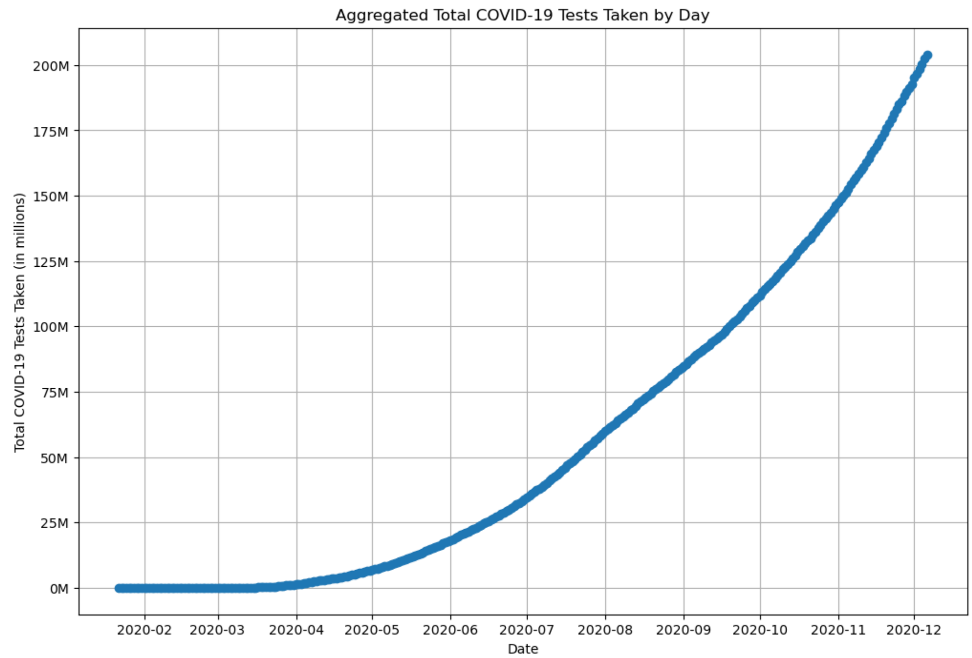

Figure 3. Total Number of Tests Taken per Day

Wrapping up the obtained statistics to get a complete picture of the evolution of COVID-19 cases over time, a good estimator would be the following figure.

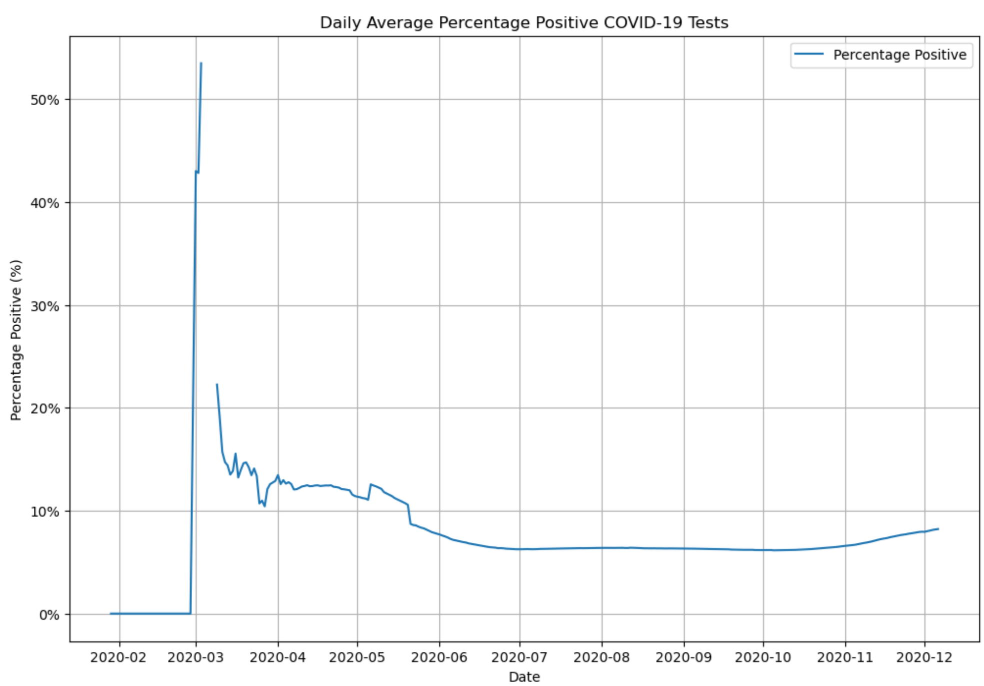

Figure 4. Share of Positive Tests Among All Taken Tests

Figure 4, which presents the percentage of positive cases relative to the total number of tests conducted, provides a more nuanced perspective on the pandemic's progression. This metric is crucial as it helps us understand the relationship between the number of tests administered and the percentage of positive results. Shows a high peak in March 2020 as the moment of greatest expansion of the pandemic in 2020. Between march and june the percentage of positive tests experienced a steady decrease that can coincide with a. Period of high awareness and prevention that caused the percentage of positive tests to drop. After June it stabilized following both the increase in number of tests taken and number of positive cases showing that they increased at similar rates.

The evolution of COVID-19 cases over time, as depicted in the earlier graphs, provides valuable insights into the virus's spread, but it is the accompanying graph showcasing the increasing deaths that serves as a poignant and tangible estimator of the true consequences of the pandemic. Deaths, representing the most severe outcomes of the virus, offer a direct measure of its impact on human lives and serve as a critical indicator for assessing the gravity of the ongoing crisis.

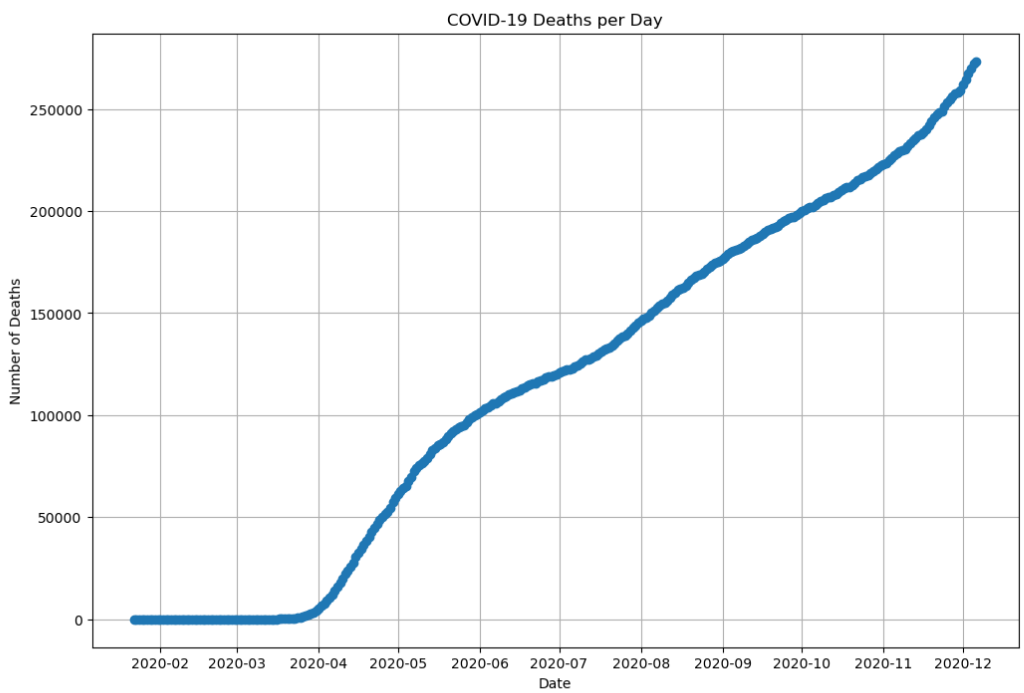

Figure 5. Total number of COVID-19 Deaths per Day

In Figure 5, we observe a steadily increasing trend in the aggregate number of COVID-19 deaths over time. This upward-sloping line represents the cumulative impact of the pandemic on mortality. The slope of this line can serve as a critical estimator, providing insights into how deaths were increasing during different phases of the pandemic, being a possible representation of the different outbreaks.

The slope of this line reveals that the pandemic's toll on human lives was relentless, with the number of deaths increasing steadily over time. It is a sobering reminder of the widespread and persistent nature of the virus. Furthermore, it highlights the importance of implementing effective public health measures and vaccination campaigns to slow the rate of increase and ultimately bring the curve under control.

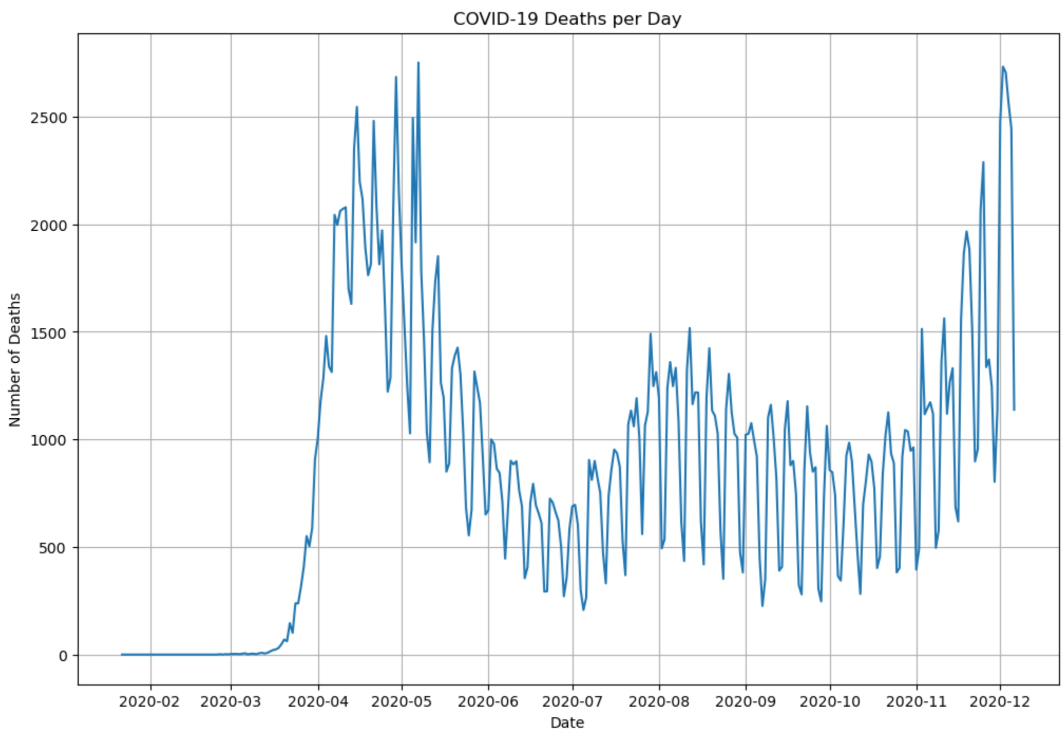

Figure 6. COVID-19 Deaths per Day

Figure 6 offers a more granular perspective on the evolution of COVID-19 deaths by depicting the daily increase in deaths over several months. The fluctuating nature of this graph reflects the dynamic interplay between the virus's outbreaks and the behaviors of individuals and communities.

From March to May, we observe a sharp increase in daily deaths, which corresponds to the initial wave of the pandemic. Subsequently, during the summer months, the graph exhibits a notable decline in deaths per day. This reduction can be attributed to increased awareness, social distancing measures, and public adherence to safety precautions during the warmer months.

However, the trend takes a different turn after the summer, as daily deaths begin to rise again. This can be attributed to a relaxation of precautions, increased social interactions, and the emergence of new outbreaks. By December, daily deaths surge significantly, reflecting the consequences of reduced vigilance and a surge in cases. This, in turn, necessitated the reimplementation of stringent quarantine measures in the following year.

In summary, Figure 6's dynamic pattern of daily COVID-19 deaths underscores the impact of behavioral changes and outbreaks on the pandemic's trajectory. It serves as a stark reminder of the importance of sustained vigilance and public health measures, especially during periods of increased social activity, to mitigate the virus's devastating effects and reduce the need for subsequent quarantine measures. This served as a great estimator of the pandemic's evolution over time.

Correlates of COVID-19: Exploring Influential Factors

To gain a deeper understanding of the dynamics of COVID-19 spread within the United States, we turn our attention to the exploration of correlation. Correlation analysis plays a pivotal role in uncovering potential relationships between various factors and the spread of the virus.

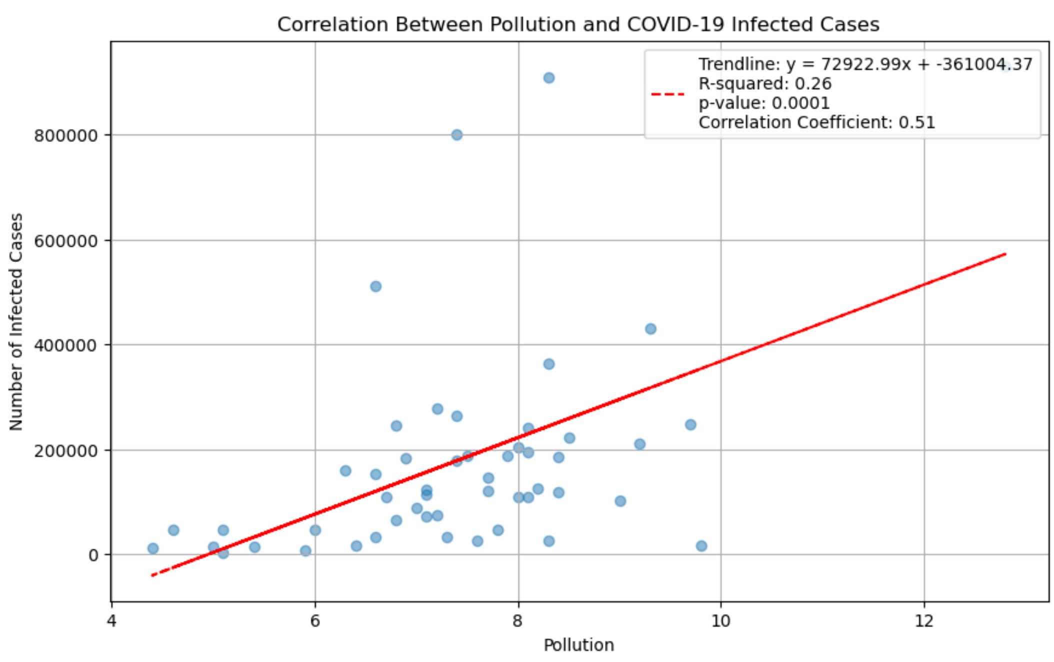

Figure 7. Correlation Between Pollution and COVID-19 Infected Cases

Figure 7 illustrates the relationship between pollution levels and the number of COVID-19 infected cases in the United States. This analysis aims to explore whether there is a significant association between air pollution and the spread of the virus. The chart's findings suggest that there is a statistically significant ($p < 0.05$) positive correlation between air pollution and the number of COVID-19 infected cases in the United States. This implies that regions with higher levels of air pollution may experience a greater number of infections. However, it's essential to acknowledge that air pollution is just one of several factors influencing the spread of the virus, and other variables, such as urban score, smoking rate, and public health measures, should also be considered in comprehensive analyses.

Having examined the correlation between pollution levels and the prevalence of COVID-19 infections, we now shift our focus to another dimension of our analysis: the correlation between urbanization and COVID-19 spread. While our exploration of pollution yielded valuable insights into environmental factors, understanding the impact of urbanization is equally pivotal. Urban environments bring their own set of dynamics, and unraveling their association with the virus's dissemination can provide critical perspectives for our comprehensive understanding.

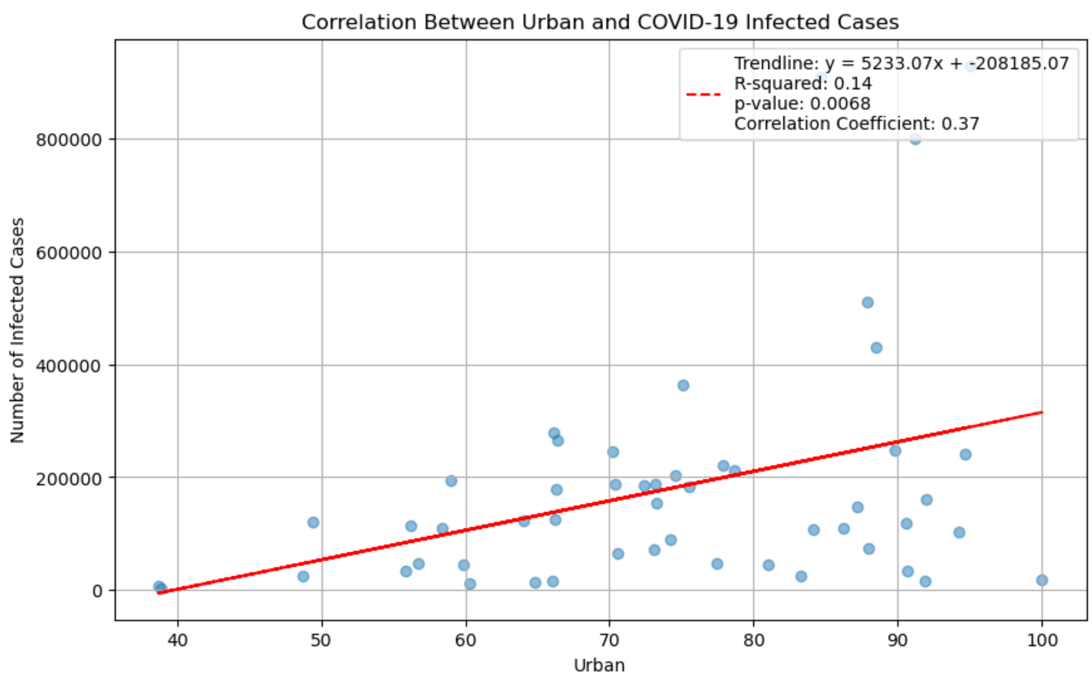

Figure 8. Correlation Between the Urban Score and COVID-19 Infected Cases

The analysis reveals that there is a statistically significant positive correlation between the degree of urbanization and the number of COVID-19 infected cases in the United States (Figure 8). This correlation suggests that regions with higher levels of urbanization are prone to experiencing a greater incidence of infections. However, this correlation does not imply causation; it underscores that urban areas may present unique conditions or dynamics that facilitate the virus's spread. Urban centers often serve as points for economic, cultural, and social activities. Consequently, they tend to have higher population densities and increased interpersonal interactions.

Continuing our exploration of factors influencing the spread and impact of COVID-19, we now turn our attention to the correlation between economic inequality, as represented by the Gini index, and the number of COVID-19 deaths (Figure 9). Earlier, we observed how urbanization levels play a role in shaping the pandemic's dynamics. Now, we delve into the socioeconomic dimension, seeking to unravel how disparities in wealth distribution may intersect with the severity of the virus's impact.

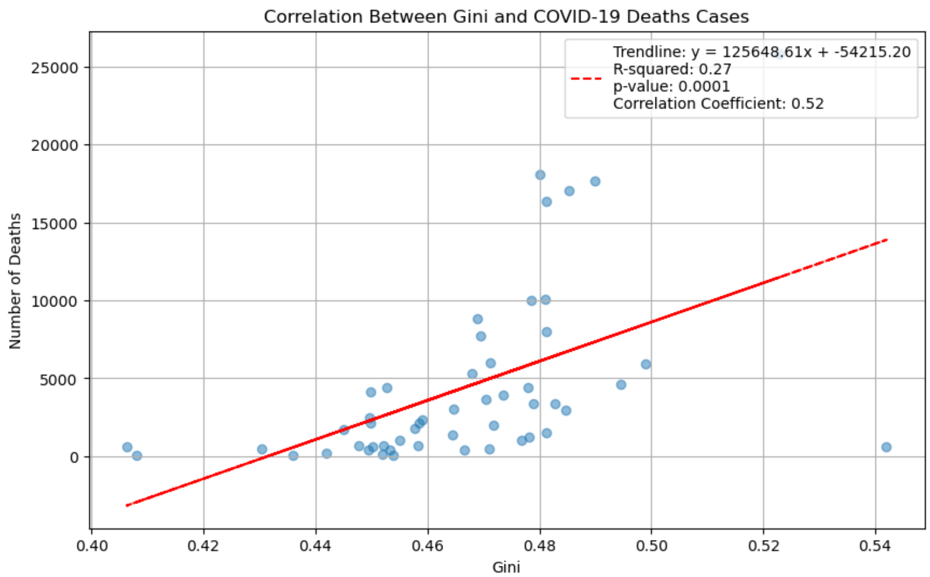

Figure 9. Correlation Between the Gini Index and COVID-19 Deaths Cases

This correlation signifies that regions characterized by higher levels of income inequality tend to experience a greater number of COVID-19 deaths. The implications of this finding are multifaceted. First, it suggests that socioeconomic disparities play a crucial role in determining the pandemic's impact, with disadvantaged populations facing a higher risk of mortality. Second, it emphasizes the need for targeted interventions and resources to address the specific vulnerabilities faced by economically disadvantaged communities. Policies that prioritize equitable access to healthcare, economic support, and vaccination resources can be instrumental in mitigating the disparate impact of the virus. Additionally, these results underscore the interconnectedness of social determinants of health and the importance of addressing systemic inequalities in our response to public health crises.

Moreover, we conducted a comprehensive analysis examining the correlation between two critical factors—ICU beds and hospitals per 100,000 population—and both COVID-19 infected cases and deaths. Our objective was to check whether the availability of healthcare resources, specifically intensive care units (ICU beds) and hospitals, had a discernible impact on the incidence of COVID-19 infections and related fatalities.

We found out that the presence of ICU beds per capita does not appear to significantly influence the number of COVID-19 cases or deaths. This suggests that factors beyond ICU bed availability, such as adherence to preventive measures and the virus's transmission dynamics, play more substantial roles. In contrast, the number of hospitals per capita exhibits a notable correlation with COVID-19 outcomes (Figure 10,11). Regions with a higher density of hospitals tend to experience fewer COVID-19 infections and related deaths. This underscores the importance of accessible healthcare infrastructure in pandemic response.

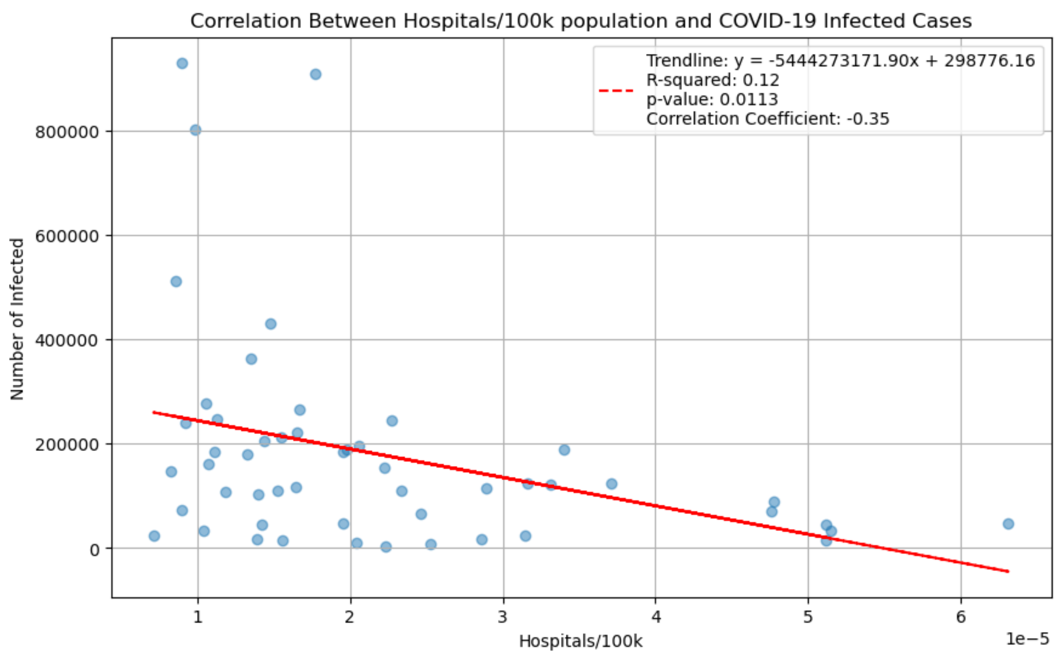

Figure 10. Correlation Between Hospitals/100k population and COVID-19 Infected Cases

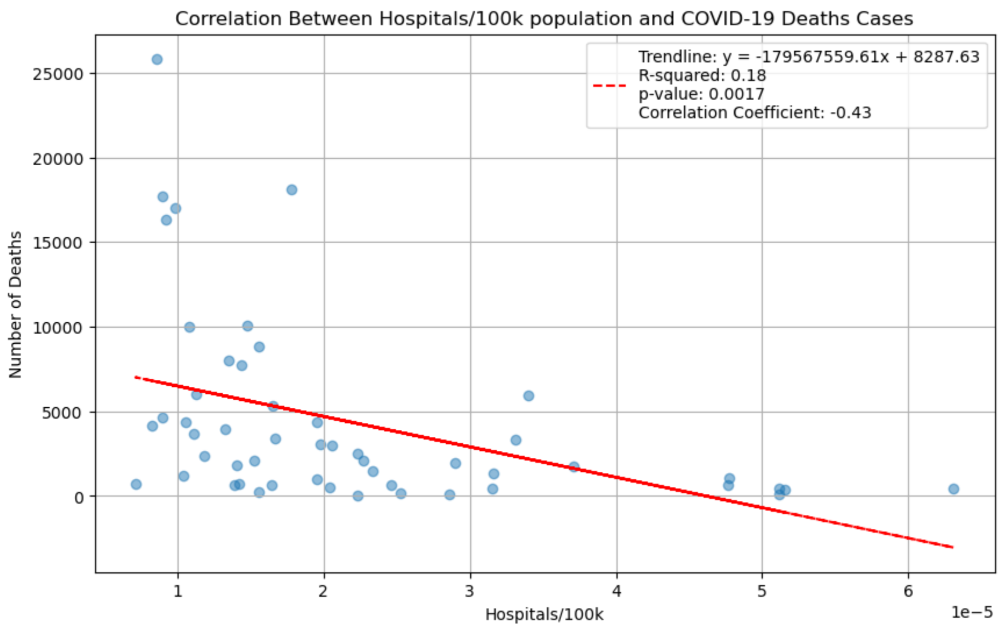

Figure 11. Correlation Between Number of Hospitals/100k population and COVID-19 Deaths Cases

In sum, our analysis underscores the multifaceted nature of the pandemic's determinants. While healthcare resources play a role, broader public health strategies encompassing prevention, testing, and vaccination remain critical components of an effective response to COVID-19.

Effect of testing

In this section, we are conducting an analysis encompassing the entire population of the United States. Consequently, all variables computed and examined represent aggregate data derived from the states' populations. This approach was chosen to gain a comprehensive understanding of the overall impact of testing and to mitigate potential biases arising from specific state-level factors that could influence the data in a similar manner as testing.

To grasp the implications of testing comprehensively, we recognize the importance of not solely focusing on absolute test counts, but also delving into the underlying dynamics. Our objective extends beyond merely assessing the presence of testing; we aim to comprehend the influence of its escalation over time.

It's crucial to elucidate why we have opted to employ daily variations instead of absolute numbers. Our investigation is centered not only on the effect of testing in isolation but also on the consequences of its incremental expansion. Utilizing these variation-based variables allows us to examine precisely how testing intensification impacts the overall pandemic landscape, which underscores the rationale for our choice in this analysis.

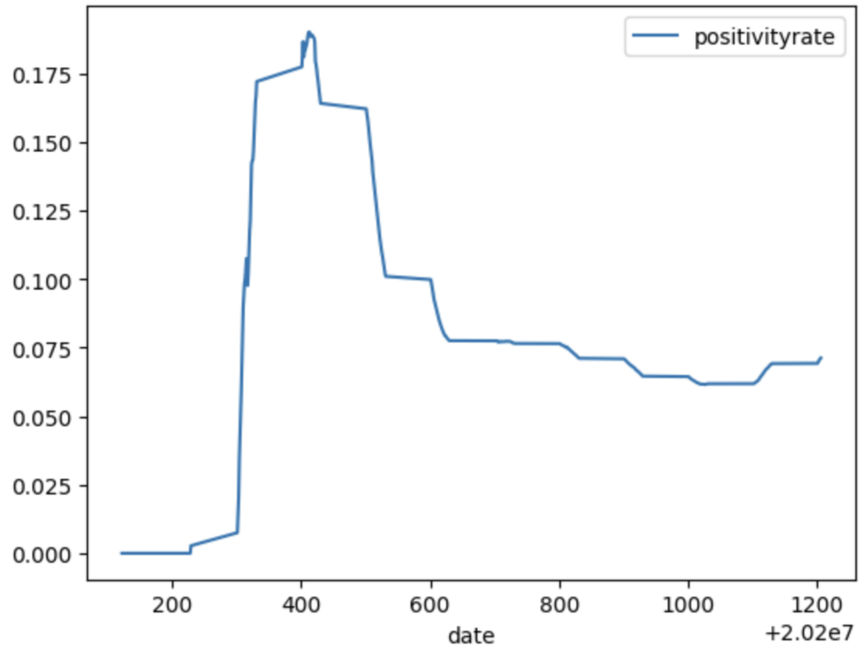

Figure 12. Positivity Rate - Ratio of positive tests and number of total tests

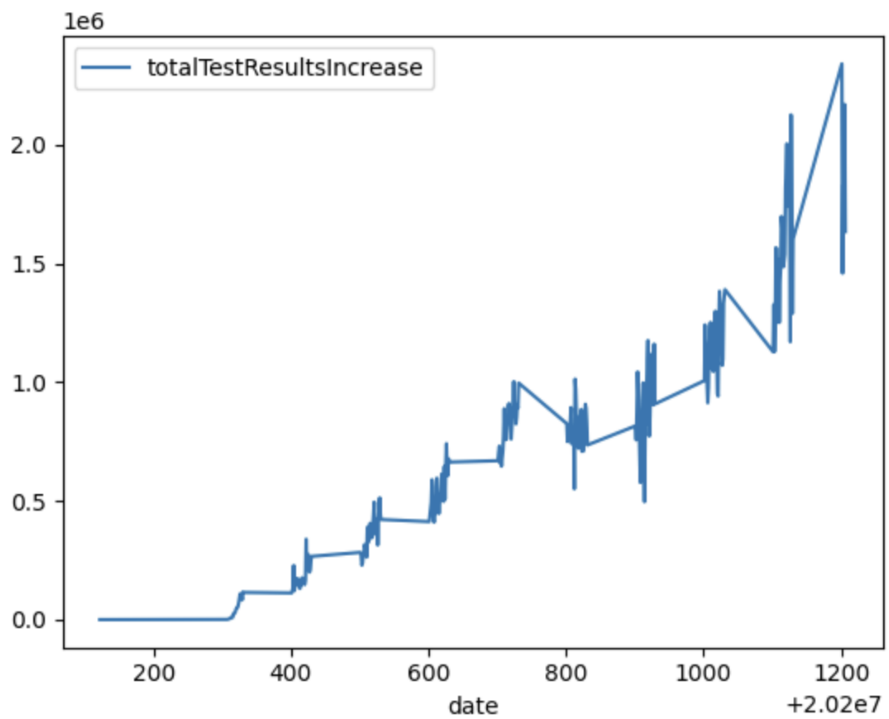

Figure 13. Daily variation of the number of the number of tests

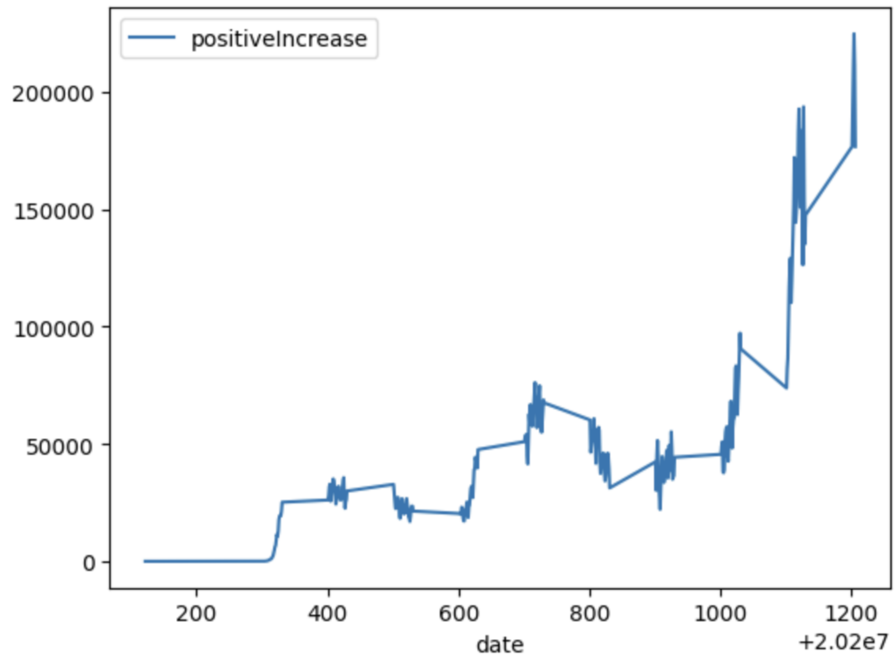

Figure 14. Daily Variation of positives per day

Considering Figure 13 and Figure 14 exclusively might lead to the misconception that conducting more tests yields a higher number of positive cases. However, delving into Figure 12 offers a distinct perspective.

Figure 12 represents the calculation of the positivity rate, achieved by dividing the number of positive cases by the total number of tests conducted. This computation is of paramount importance as it provides insight into potential sample biases in the early stages of testing. It suggests that initially, when testing was limited, those who underwent testing were predominantly individuals displaying symptoms. Alternatively, this observation could hint at a rapid initial spread followed by a subsequent decrease. While these are hypotheses, what we can ascertain is that the widespread availability of tests has led to the stabilization of the positivity rate.

Furthermore, it is evident that testing became more widespread as the total number of tests substantially increased. Nevertheless, to conclusively determine if the positivity rate has indeed stabilized, it becomes imperative to gauge what proportion of the overall population has been tested in recent days. Without this data, our conclusions remain somewhat speculative, considering the potential sample bias inherent in those who seek testing due to symptomatic presentation.

To gain a more precise understanding of the testing's impact, it is essential to examine data related to infected people, hospitalizations and fatalities caused by severe COVID-19 cases among those who undergo testing.

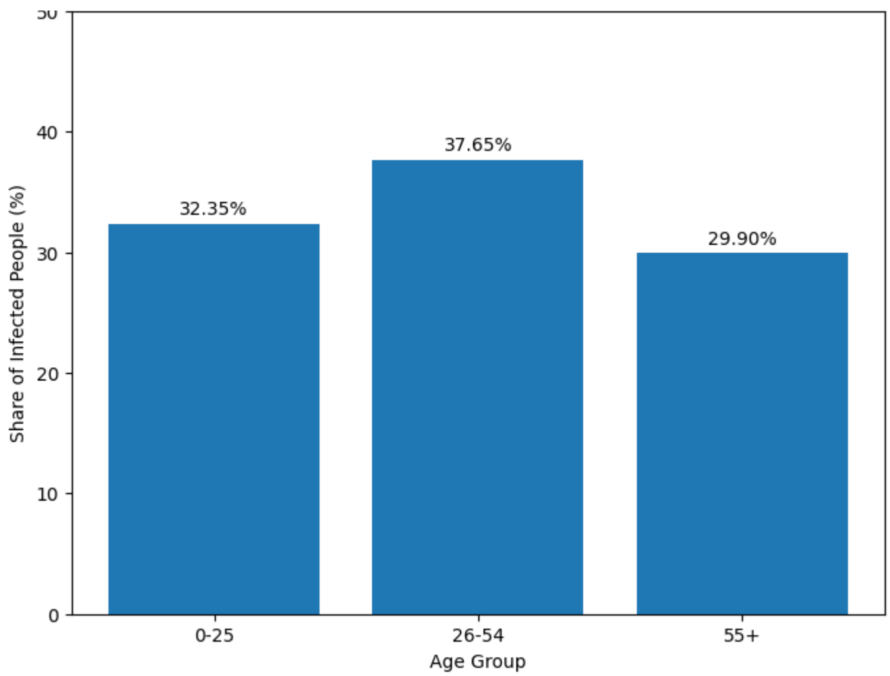

Figure 15. Share of infected people in each age group

Figure 15 illustrates the share of infected people (%), categorized by age groups where we can see that the subgroups that suffered the most are as follow:

Age group 26-54: This age group has the highest share of infected individuals at 37.65%. This suggests that people between the ages of 26 and 54 have been significantly impacted by the infection.

Age group 0-25: The second-highest share of infected people falls within the age group 0-25, with a share of 32.35%. While slightly lower than the 26-54 age group, this still represents a substantial portion of the infected population, particularly affecting children, teenagers, and young adults.

Age group 55+: The age group 55+ has the lowest share of infected individuals at 29.90%. While this group has the lowest percentage of infections among the three groups, it's important to note that older individuals in this group may be at a higher risk of severe illness or complications from the infection, despite having a lower infection rate.

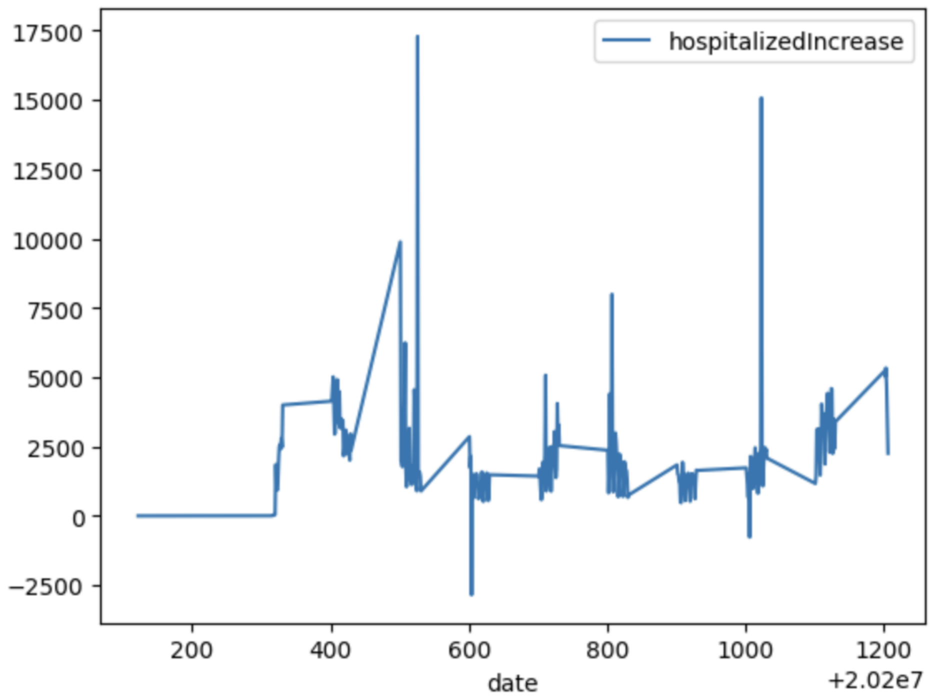

Figure 16. Daily Variation of hospitalizations per day

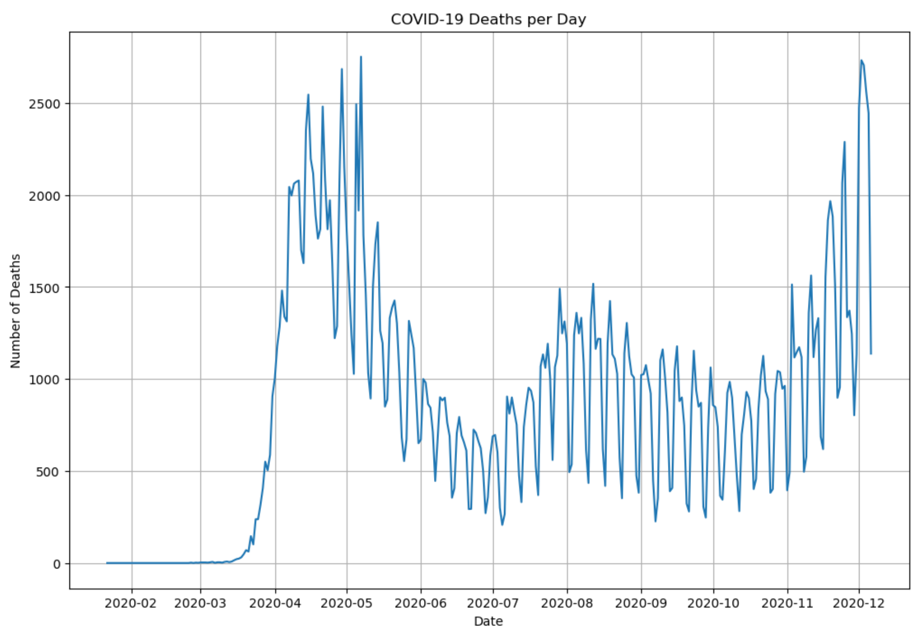

Figure 17. Daily Variation of deaths per day

Upon scrutinizing this dataset, a compelling pattern emerges: in the early stages, a rise in the positivity rate coincides with an uptick in daily hospitalizations and fatalities. Assuming the disease's nature remains unchanged, this suggests that individuals who sought testing in the initial phase exhibited more severe symptoms, resulting in a corresponding increase in deaths and hospital admissions.

Conversely, we observe that the trajectories of testing and positive cases do not mirror the dynamics of COVID-19's effects. This reinforces our inference that an escalation in testing does not necessarily correspond to an increase in hospitalizations and deaths.

Hence, the calculation of the positivity rate assumes significant importance. When combined with other data, it offers valuable insights—while testing may uncover more cases and positives, the unaltered proportion suggests that milder cases are also being detected.

In the latter part of the dataset, we note an upswing in both deaths and hospitalizations, accompanied by a modest increase in the positivity rate and positive cases. While this may hint at a potential shift, the available data does not provide sufficient information for conclusive inferences.

In summary, this dataset underscores the importance of widespread testing to gain a comprehensive understanding of COVID-19 dynamics. Initially, a robust correlation existed between the positivity rate and both fatalities and hospitalizations. However, in the latter part of the dataset, this correlation diminishes. Further examination reveals a significant increase in the number of tests conducted (exceeding positive tests), offering a more comprehensive perspective of the pandemic's reality. This emphasizes the need for expansive testing to better comprehend the true landscape of COVID-19 and prepare for any significant fluctuations in positive cases, particularly the more severe manifestations of the disease.

Contagion Model and Policy Simulation

For better understanding of COVID-19 spread, we implemented a mathematical contagion model. Through this simulation, we aim to provide valuable insights that can guide policymakers and healthcare professionals in the United States and beyond.

Methodology:

- Model Description: We implemented a mathematical contagion model based on a population of 1000 people. Each individual has the potential to infect up to 3 others per day with a probability of 0.05.
- Initial Infections: To initiate the simulation, we assumed that 10 individuals were initially infected with COVID-19.

In the short term, our simulation revealed that after 3 days, the number of infected individuals increased to 15 (Figure 18). This outcome underscores the rapid spread of the virus, even with a relatively low transmission probability.

Furthermore, our long-term projection (Figure 19) highlighted a sobering reality: it is expected to take approximately 33 days for the entire population to be infected, emphasizing the need for early and sustained intervention efforts. These findings underscore the critical importance of immediate and comprehensive public health measures, including vaccination campaigns, social distancing, mask mandates, and adaptive policymaking, to mitigate the virus's impact and protect our communities.

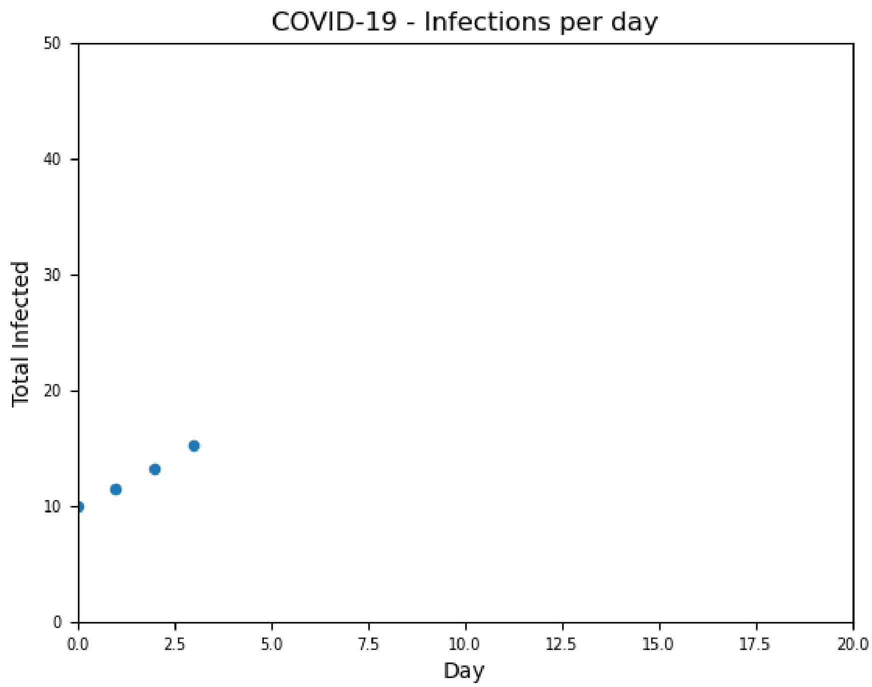

Figure 18. Number of infected people after 3 days

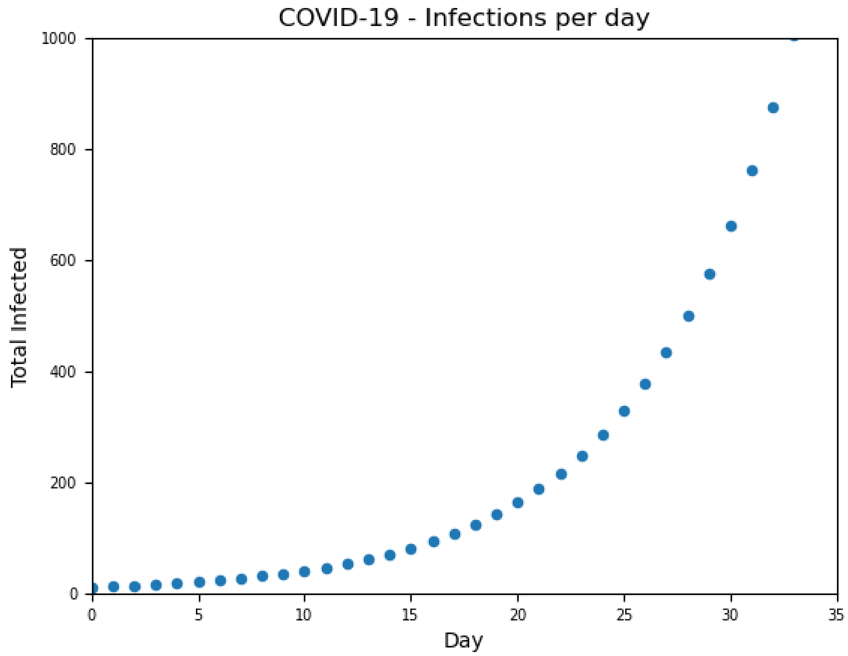

Figure 19. Number of people infected per day (till the full population is infected)

In our simulation study focused on combating the spread of COVID-19, we implemented a vaccination policy on day 3 (Figure 20), introducing immunity to uninfected individuals and reducing their infection probability by 0.01 percentage points. The results were striking: after just 10 days, the number of infections decreased significantly to approximately 19 individuals, showcasing the effectiveness of early vaccination in curbing the virus's spread. Furthermore, our findings demonstrated that introducing vaccination could extend the time to full population infection to approximately 145 days, providing crucial time for healthcare system preparedness. These outcomes underscore the pivotal role of vaccination as a cornerstone in public health strategies, offering both short-term relief and long-term pandemic management potential.

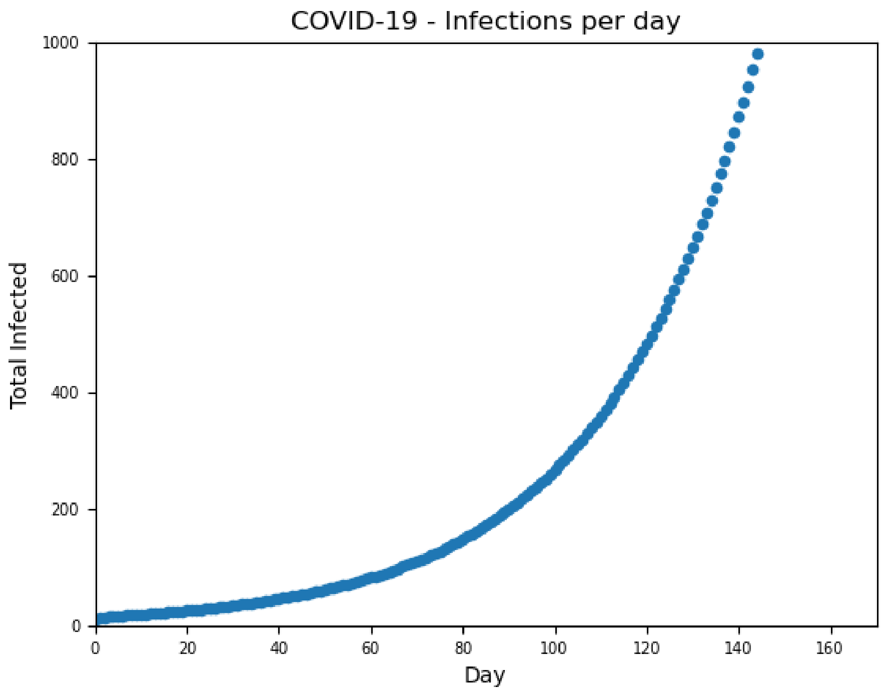

Figure 20. Number of COVID-19 infections per day given that vaccination policy of 3 days was introduced

In conclusion, our mathematical contagion model and subsequent vaccination policy simulation underscore the significance of timely interventions in combating COVID-19. The model revealed the virus's rapid spread without intervention, but the vaccination policy introduced on day 3 notably reduced infections by day 10 and extended the timeline for full population infection to approximately 145 days. These findings emphasize the critical role of vaccination in flattening the curve, protecting vulnerable groups, and providing essential time for healthcare system preparedness. As the pandemic continues, our study reinforces the urgency of vaccination campaigns and the importance of ongoing vigilance in managing COVID-19 effectively.

Recommendations

Based on the performed analysis, here are recommendations for targeting measures to reduce the spread of COVID-19:

1. **Consider Urban Areas**: The correlation analysis revealed a positive association between urbanization and the number of COVID-19 infected cases. Thus, urban areas should be a critical point for confinement measures. Implement measures such as social distancing, mask mandates, and restrictions on public gatherings in densely populated urban regions to curb transmission.

2. **Address Socioeconomic Disparities**: The correlation between the Gini index (income inequality) and COVID-19 deaths indicates that economically disadvantaged populations may be more vulnerable. Thus, economic support, access to healthcare, and vaccination resources should be provided to disadvantaged communities.
3. **Enhance Healthcare Infrastructure**: The presence of hospitals per capita showed a negative correlation with both COVID-19 infections and deaths. Prioritize regions with fewer healthcare facilities for resource allocation and support. Ensure that healthcare infrastructure is robust and accessible to provide timely care and treatment.
4. **Comprehensive Strategies**: Recognize that the pandemic's dynamics are influenced by a multitude of factors beyond those explored in correlations. Comprehensive strategies must encompass prevention, testing, contact tracing, vaccination, and public health messaging. These strategies should be adaptable and responsive to evolving circumstances.
5. **Vulnerable Groups**: Identify and prioritize vulnerable populations, including the elderly, individuals with underlying health conditions, and essential workers, for vaccination and targeted protection measures. Protecting these groups can significantly reduce severe outcomes.
6. **Adapt and Evolve:** Recognize that the pandemic is dynamic, and confinement strategies should adapt as the situation evolves. Flexibility in response strategies is essential to effectively manage the contagion.
7. **Test more:** To understand the dynamics of covid in the overall population, the widespread of testing can lead to capture not only people with symptoms, but also people without any symptoms. As these people also transmit covid, it is very important to catch more positives to further anticipate any increase in the more harmful cases of covid.

These recommendations underscore the importance of a targeted and evidence-based approach to confinement measures. By focusing on specific subgroups and regions identified through correlation analysis, authorities can optimize resource allocation and tailor interventions to mitigate the spread of COVID-19 and protect vulnerable populations.

References:

Davahli, M. R., Karwowski, W., Sonmez, S., & Apostolopoulos, Y. (2020). The hospitality industry in the face of the COVID-19 pandemic: Current topics and research methods. *International journal of environmental research and public health*, *17*(20), 7366.

Gursoy, D., & Chi, C. G. (2020). Effects of COVID-19 pandemic on hospitality industry: review of the current situations and a research agenda. *Journal of Hospitality Marketing & Management*, *29*(5), 527-529.

OECD. 2020. *"The territorial impact of COVID-19: Managing the crisis across levels of government"*. Retrived from https://www.oecd.org/coronavirus/policy-responses/the-territorial-impact-of-covid-19-managing-the-crisis-across-levels-of-government-d3e314e1/

CEPAL. 2020. "*Impact of COVID-19 on the United States economy and the policy response"*. Retrieved from https://repositorio.cepal.org/server/api/core/bitstreams/9adde514-a4f4-494d-9aa1-8ecd5712158b/content