Navigating Pandemic Realities: A Data Visualization Exploration

Jiaqin Wu

2023-11-23

Executive Summary

Amid the extraordinary challenges presented by the COVID-19 pandemic, this initiative stands as a guiding light, venturing into the complexities of public-health issues throughout the United States. Leveraging the comprehensive dataset from the U.S. Department of Health & Human Services, we have meticulously translated raw daily data into a refined monthly format, carefully curating the most insightful variables for a thorough exploration. Our journey unfolds through an innovative and user-friendly Shiny app, crafted to unveil the nuanced dimensions of the pandemic's impact.

Comprising four distinct panels, our exploration traverses through the maze of information, providing profound insights into the distribution of confirmed and suspected cases over time and by region, age-specific case distributions over time and by region, dynamics of inpatient bed utilization over time and by region, and the geographic distribution of death cases across diverse states over time. Anchored by Washington, DC as the baseline, this report narrates the public health policy story in the nation's capital, offering a panoramic view of the contextual backdrop. Furthermore, audiences can delve into the conditions of other states by interacting with various buttons within the app. In the Shiny app, our narrative extends beyond the scope of COVID situations in Washington, DC, providing an overview of selected indicators from the dataset that depict the overall state of public health.

Quick Links

- Link to the original dataset: HeathData.gov
- Link to the visualization: Shiny Application
- Link to the codebooks: GitHub

Current Context

In the current context, the project addresses the pressing challenges posed by the ongoing COVID-19 pandemic, a global health crisis that has significantly impacted communities worldwide. The project's focus lies in unraveling the intricacies of public health issues across the United States, leveraging a comprehensive dataset from the U.S. Department of Health & Human Services. As the pandemic continues to evolve, there is a critical need for in-depth exploration and analysis of health-related data to guide effective decision-making and interventions. The project provides a timely response to the heightened demand for accessible, user-friendly tools that offer insights into the distribution of confirmed and suspected cases, dynamics of healthcare resource utilization, and the geographic spread of mortality. By delving into these critical aspects, the project contributes to a better understanding of the pandemic's impact, ultimately supporting informed policymaking, resource allocation, and public health strategies.

Intended Audiences

The project caters to a broad audience with a shared interest in addressing the complex challenges posed by the COVID-19 pandemic through data-driven insights. The intended audience of this project encompasses a diverse range of stakeholders involved in public health, policy, and data analysis. Primarily, healthcare professionals and practitioners stand to benefit from the insights derived through the comprehensive exploration of COVID-19 data, aiding them in making informed decisions about resource allocation, patient care, and preventive measures. Policymakers at various levels of government can leverage the findings to formulate effective strategies and interventions. Additionally, researchers and academics focused on epidemiology and public health will find value in the detailed analysis and visualizations provided by the project. The general public, seeking a nuanced understanding of the pandemic's impact on different regions, demographics, and healthcare resources, can also benefit from the accessible and user-friendly presentation of information.

Data Sources and Transformations

- **Source:** U.S. Department of Health & Human Services
- **License:** N/A
- **Collection Methods:** The U.S. Department of Health & Human Services meticulously gathered data through the everyday reports submitted by hospitals across different states, creating a comprehensive snapshot of the evolving public-health landscape during the pandemic.

- **Biases/Sampling:** In the realm of public health data, the potential for biases stemming from variations in reported hospitalization numbers looms as a significant challenge. Acknowledging this complexity, our analytical approach adopts a nuanced strategy to mitigate potential biases. For example, when comparing the proportion of inpatient beds used for COVID, variations in reporting across hospitals introduce a layer of intricacy. The disparities in reported figures necessitate a meticulous examination to ensure the comparability of variables, highlighting the importance of a cautious and informed analysis.
- **Data Cleaning Methods & Identified Issues:** The original dataset, brimming with variables, prompted us to discern the most insightful among them. Our meticulous selection and transformation of these variables align with our pursuit of meaningful insights. Shaping the dataset to our objectives involved the transformation of day-level variables into a more manageable month-level format. Variables like "covid_confirmed" and "covid_suspected" underwent a metamorphosis into "value," "coverage," and "COVID," enriching our data with information about the COVID status and whether the data pertains to case numbers or reported hospitalizations. Similarly, variables reflecting demographic and age groups underwent transformation, contributing to a final dataset with lower dimensionality that balances informativeness and conciseness.
- **Descriptive Statistics:** Navigating the transition from day-level to month-level data, our commitment to accuracy led us to omit incomplete data for November 2023. Our final dataset now spans from March 2020 to October 2023. In refining variables, we reintroduced columns such as age bands and demographic groups. Demographic groups were judiciously classified into "adult" and "pediatric," encompassing individuals over and under 18, respectively. State information comprehensively covers the entire United States, and the date information is meticulously structured to align with existing months in the dataset for each year (2020-2023). This rigorous process ensures the integrity and relevance of our final datasets, empowering our analyses with clarity and depth.
- **Privacy:** Public access is allowed.

Technologies/Platforms Used

- Shiny (R): Our chosen powerhouse for crafting an interactive and user-friendly dashboard, enabling seamless exploration of the visualized data.
- RStudio: The coding arena where we brought our visualization functions to life, creating a dynamic and engaging user experience.
- Python: Our trusted ally in the data cleaning process, ensuring the integrity and accuracy of our datasets. With Python, we navigated through the intricacies of data transformation to present you with refined and meaningful insights.

Key Insights

The core objective of this visualization project is to offer an interactive and insightful platform for comprehending the profound impact of COVID-19 on public health issues. Our exploration unfolds across key dimensions:

- **Case Numbers Variation:** We delve into the fluctuations of case numbers across states and age bands, scrutinizing whether specific demographics or regions exhibit higher infection rates. This nuanced analysis provides valuable insights for local governments to implement targeted prevention measures, anticipating and mitigating potential public health challenges.
- **Inpatient Bed Utilization:** By examining the total number of beds, bed usage rates, and the proportion of beds allocated for COVID treatment, we aim to uncover patterns indicative of strained healthcare resources. This exploration not only highlights critical periods of high inpatient bed utilization but also informs strategic decisions, such as allocating funds for additional hospital infrastructure to fortify our healthcare system against future emergencies.
- **Geographic Visualization of Death Cases:** Our spatial representation of death cases across time and states acts as a dynamic tool for identifying urgent areas and time periods that demand heightened attention and resources. This visualization aids policymakers in directing funds to states or specific periods with pressing public health needs.
- **Horizontal and Vertical Comparisons:** We introduce the number of hospitalizations as a critical factor for both horizontal and vertical comparisons. This detailed analysis allows for robust comparisons between different states and time periods, enhancing the credibility of our conclusions and facilitating more convincing public health decisions.
- **Shiny App Functionality:** To empower users, our Shiny app enables the exploration and download of preprocessed datasets and visualizations in different states and time periods. Users can selectively choose variables for comparison, fostering transparency and facilitating deeper analysis. This feature enhances the accessibility of our findings, encouraging a wider audience to engage with and interpret the data.

Recommendations:

- Consider targeted preventive measures in demographics and regions with higher infection rates.
- Evaluate the need for additional healthcare infrastructure during critical periods in some states with high inpatient bed utilization rate for COVID.
- Direct funds to regions and time periods highlighted by spatial distribution of death cases.

Summary of Analysis

Our in-depth analysis focuses on pivotal dimensions of the impact of COVID-19 on public health. By scrutinizing case numbers, bed utilization, and the geographical distribution of death cases, we aim to offer precise and actionable insights. These insights serve as valuable tools for guiding preventive measures, optimizing resource allocation, and facilitating informed decision-making. The accessibility of our user-friendly Shiny app ensures that users can effortlessly navigate the dashboard, making it a powerful tool for gaining valuable insights into the complexities of the ongoing public health crisis.

Notes

For more information, please visit the GitHub page!