Impact of the Covid-19 epidemic on the U.S. economy

Yanli Wang (qkn122) Feng Feng (kvh 423) Michael Widdowson (klh272)

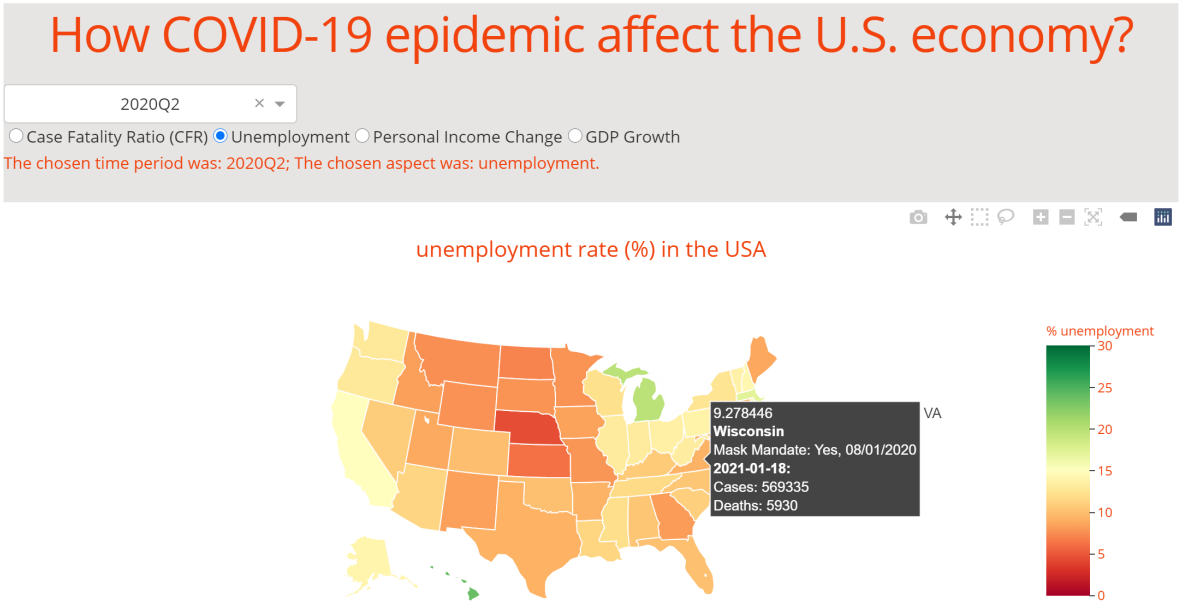

Fig. 1. Screenshot of interactive visualization: based on the U.S economy data by quarter (2019Q3 - 2020Q1), contains unemployment, personal income change, GDP growth aspects.

Abstract—How does the COVID-19 epidemic affect the global economy? In order to explore this question, we chose the U.S. as our sample case for analysis due to its abundance of data - our visualization results should also be applicable to other economies or countries. Here we showed the impact of COVID-19 on the U.S. economy in several aspects, including GDP growth, personal income changes, and unemployment rate, which ranges from the third fiscal quarter (2019Q3) to present (2021Q1) for comparison. Our visualization results also contain the changes in the case fatality rate (CFR) during the course of the epidemic, which can be used as a comparative observation with the economy. The visualization mainly uses a choropleth graph as its backbone, a drop-down menu and radio-items as the selection components, and a color bar to indicate the rate of change (percentage) for the different aspects. Our target viewers are all people who are concerned about the epidemic situation and its economic impact, especially those who work on the people's employment and livelihood.

1 Introduction

As observed in the economic outcome of 2020 (Appendix Fig.1), the global economy suffered greatly due to the havoc caused by the COVID-19 pandemic. With the exception of China - albeit, one of the slowest economic growths China has seen over the last couple of years - just about every country's economy in the world experienced a negative impact (International Monetary Fund, 2021). In this report, we wanted to explore the data available for 2020 on a macro and micro level, specifically the United States of America, to see what findings we could uncover on the different scales to tell a deeper story of the past year's crisis. The United States was chosen to be scrutinized on a more local level primarily because of the richness of COVID and non-COVID datasets available, as well as a large number of widespread cases - making it a unique example among modernized nations. It would have been unfeasible to consider every country that had a negative outcome for 2020 microscopically, so we stuck with what we believed to be abundant and relevant - being that the United States plays a central role in the global economy. Here in this report, we started with global data and told a story of what happened in the United States with regards to the COVID-19 pandemic.

Because economics is a quite complex subject (we are not economists ourselves) and many factors have influence over it, we decided to focus on what we saw as important, or rather interesting, facets to consider when trying to probe the global economic downfall. There is no doubt that a pandemic would have an impact on the global economy. And because of the rich data resources, we will explore the United States in more detail with regards to economic measurements and its unique COVID-19 data. What impact did the number of COVID-19 cases have on a country's GDP? What about deaths? Did the human development index (HDI) play a role? In the United States, what impact did mask mandates have? Did these results synthesize with each other to cause the country's outcome? How did the United States' unemployment rate transpire over the course of the epidemic? Could we observe anything from Personal Income Change? These were just a

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

few aspects we considered when approaching these inquiries worldwide and in the United States.

This topic was of particular interest to us because we wanted to shed light on the relationship between covid-19 and economy. Moreover, nearly every individual has been affected in some way by this pandemic, whether directly or indirectly; therefore, we decided to elucidate what factors we thought might influence these outcomes in hopes of providing some insight into what may have contributed to this outcome.

1.1 Related work

Economists generally believe that the Covid-19 epidemic will have a serious negative impact on the global economy. Due to lockdown measures and social distancing, the economic losses caused by the COVID-19 pandemic are largely driven by the decline in demand. This change can be clearly seen in severely affected industries such as travel and tourism. In addition, the fact is that once the pandemic is over, the global economy is at least likely to experience a sharp rebound.

For the United States, the pandemic had destroyed people's lives, brought the hospital system to its limits, and caused a public health crisis in addition to an economic crisis. Up until now, due to the improvement of the epidemic policy and the development of vaccines, the US economy has turned from the weakness in 2020 to signs of recovery. Many media and research departments have conducted visualization and analysis on various aspects of data during the epidemic so that people can understand the current status of the impact of Covid-19. For example, the Back-to-Normal Index allows people to understand how much economic activity has returned to the level before the epidemic in March (Tracking America's recovery, 2021). There are many other indicators, such as unemployment claims, getting hired (based on the number of LinkedIn users who added a new employer to their profile during this period), linkin job postings, household expenditures, personal savings, etc. We decided to delve into some of these aspects ourselves, and to include aspects of our own, for our visualizations.

1.2 Design goal

Our main goal was to analyze the COVID-19 data of the United States and a series of economic-development related data, and to explore its impact on the US economy. Our target audience are those who want to know about the development of the COVID-19 epidemic in the United States and its related impact on the economy. We started our research from global data and selected the United States as a key research object, and visualized it from multiple aspects. The main visualization plot contains various data for different quarters from 2020Q1-2021Q1, including case fatality rate (CFR) representing a measure of disease severity (Harrington, 2016), as well as various economic-development related data such as GDP growth, unemployment rate, personal income change, etc. This allowed us to create a visual representation of the impact the COVID-19 pandemic had on the United States over the past year.

With the combination of these distinct datasets we created unique visualizations not found anywhere else. The full interactivity of them allows the user to explore the data however they please. Our simplification of complex data allows for an intuitive experience of the data. Because the data is time-series based, the user can observe trends across time with respect to cases, deaths, case fatality rate, mask mandates, unemployment rates, GDP growth, and personal income changes. We also wanted to keep the coding behind these visualizations simple enough so that it can be redeployed, or repurposed, on other nations that possess that necessary data. It was an underlying goal to bring complexity to elegant simplicity, so that our main goal in question could be as clear and straight-forward as possible for our target audience.

2 Methodology

2.1 Global COVID-19 and economy

The global dataset is from a Kaggle dataset called Impact of Covid-19 Pandemic on the Global Economy (Tiwari, 2020), which contains timestamps, total cases, total deaths, stringency index, population, gdp per capita, and human development index - 8 features in total from 210 countries or regions from 2019-12 to 2020-10.

Location: 210 countries or regions.

Stringency Index (Coronavirus Government Response Tracker, 2020): Every country has different lockdown policies, and stringency index measures the strictness of lockdown. This is a constructed composite index based on the simple average of nine response indicators including school closures, workplace closures, travel bans, etc. The value is rescaled from 0 to 100 (100 = strictest). (min=0, mean=48.135, max=91.332)

Human Development Index (HDI; Human Development Reports, n.d.): The HDI is a composite statistics index based on income per capita, education and life expectancy. It aims to measure the development degree of a country, not only economy or GDP growth, as two countries with similar GDP per capita level in their economy could lead to different living standards of people. The value ranges from 0-1.0 (1.0 highest). (min=0, mean=0.612, max=0.953)

GDP per capita: The gross domestic product per capita (GDP per capita) divides the country's gross domestic product by its total population. (min=0, mean=16,991.085, max=116,935.600)

Case/pop: We compute this metric by the confirmed COVID-19 cases divided by the population in a country, which measures the probability of COVID-19 confirmed in the population. The value ranges from 0 to 1. (min=0, mean=0.00216, max=0.0193)

First of all, we used the correlation diagram, an interactive scatter plot about GDP, HDI, control, and COVID-19, to explore the relationship between multiple economic factors and the COVID-19 epidemic. Through the graph, it's hard to find the obvious connection between COVID-19 cases and the economy, thus we transform the raw data by computing cases/pop, deaths/pop, and deaths/cases, etc. With that that we found GDP per capita is somewhat related to HDI (owed to its definition) and the probability of COVID-19 confirmed in the population. Then we plotted these features into one scatter graph to show the linear relationship between them, where the GDP per capita is transformed into logarithmic scale.

The scatter plot is mainly composed of 210 scatters, where each point represents an economic entity, and the size of a scatter point represents cases/pop. The larger the circle, the higher the confirmed rate of COVID-19 among the local population; the x-axis represents the human development index, the higher the better people's development; the vertical axis represents log (GDP per capita); the color of each scatter point represents the stringency index, and the right-side color bar maps the index into a color range, the higher value (tends to yellow) means more strictness of lockdown. In addition, adding death/case information in hover information to compare with cases/pop.

2.2 Monthly new cases/death animation Of the world and the US

This part is based on the COVID-19 dataset from John Hopkins University (Goldbloom, 2020). This dataset has new cases and deaths data for all countries in the world and also all counties in the US, which is suitable for making animation.

We got the daily new global data from the "CONVENIENT_global_confirmed_cases.csv" and "CONVENIENT_global_deaths.csv", which includes 191 countries and regions, from 23/1/2020 to 17/1/2021. The range of cases is from -47301 to 823225], the range of deaths is from -1918 to 4462].

We changed all negative values to zero, transformed the data to long format and calculated the monthly cases or deaths. After data cleaning, we made the animations of choropleth plots with the selection for "cases" or "deaths" by means of plotly express.

Next, we analyzed the U.S. epidemic in more depth. We used the daily new USA data from "CONVENIENT_us_confirmed_cases.csv" and "CONVENIENT_us_deaths.csv" , which includes 3340 counties and 58 states/regions, from 23/1/2020 to 17/1/2021.The range of cases is [-7590, 29423], the range of deaths is [-3962,1553]. We did similar data processing, and added location information from "RAW_us_confirmed_cases.csv". We then removed some special locations without latitude and longitude, like Diamond Princess and Grand Princess. After data cleaning, we finally got the monthly data of 52 states and 3228 countries/regions, and then made the animations of

scatter plots with the selection for “cases” or “deaths”, where the size or color both represent the monthly new cases or death.

2.3 Dash plot

Apart from the illustration of COVID-19 cases and deaths on the map, we also want to detect whether there exists some underlying relationship between the epidemic and the economy - this is also our main goal of visualization. In this part, we merged a few economic data and the COVID-19 cases(deaths) data in the U.S., mainly focusing on unemployment, personal income, and GDP growth rate. We believe these three factors can describe the micro-and macro-economy in the U.S. to some extent.

Qtr: time period (quarter) , from 2019Q3 to 2021Q1p(2021-01-18); Cases: cumulative confirmed COVID-19 cases quarterly (the max value in a quarter) Deaths: cumulative deaths caused by COVID-19 quarterly (the max value in a quarter) Case Fatality Ratio (CFR): estimates the proportion of deaths among confirmed COVID-19 cases quarterly (the mean value in a quarter) Unemployment Rate: the percentage rate of unemployment (min=2.189, mean=6.598, max=23.867) Personal Income Change: the percentage change rate of personal income, including the income that people get from salaries, government benefits, dividends and business interest and other sources. (min=-29.900, mean=7.258, max=70.300) GDP growth rate: the rate of GDP growth; measure of U.S. economic activity (min=-42.200, mean=0.513, max=52.200) Cases_p: cumulative confirmed COVID-19 cases till now (2021-01-18) Deaths_p: cumulative confirmed COVID-19 cases till now (2021-01-18) CFR_p: estimates the proportion of deaths among confirmed COVID-19 cases (2021-01-18) (min=0.400, mean=1.581, max=3.200) Mask: mask mandate (Yes or No) and if Yes, the start date.

Since the economic data are mostly recorded quarterly, we processed the merged dataset (including unemployment, personal income, GDP, mask, and COVID-19 data), put all of them into quarterly timestamps, and filled missing places with np.NaN values. Besides, the datasets originate from different sources, spanning different lengths of time period, we thus set our observations from 2019Q3 to 2021Q1.

We built up a framework to visualize our merged data. It mainly contains three parts: the header with filtering selection menus - the dropdown menu to select a single time period (quarter) and radio items indicates which economic aspects or the COVID-19 CFR; the choropleth graph colored by the values of selected aspect, in which hovered information about the name of State and the COVID-19 status at present (2021-01-18); the right-side color-bar indicates the range of values scale, for the same aspect, which we fixed the color mapped range (max and min) among all the quarters. Some reminder texts are also included: when the user has selected the time period (default to ”Select a time period (quarter)” as None) and the aspect selected on the radio items (default to ”Unemployment”), there shows a line to help the user double-check. The graph title also changes with the selected aspect, likewise, the title of the color-bar indicates the same information, which reduces the misleading information during the conveying process. We especially tried a variety of color-scales and finally decided to set the color-bar in the range of red to green, which is also commonly used in the economic field, and yellow represents neutral (value=0), which is differentiated from white that indicates missing data (np.NaN). In addition, the hover texts also show the mask mandate status in each state.

3 Result

3.1 Global

It is very easy to observe the linear relationship between the human development index and logarithmic GDP per capita through the scatter plot, except for a very small number of countries that clustered into outliers. Meanwhile, we also observe that countries with better development and economic levels tend to have a higher case/pop ratio (the larger diameter of the scatter), which seems contrary to common sense that we subconsciously think that developed countries have better epidemic control. We guess this is attributed to the much more frequent flow in developed economy entities than those under-developed countries. The stringency index (scatter color) can tell us how strict the

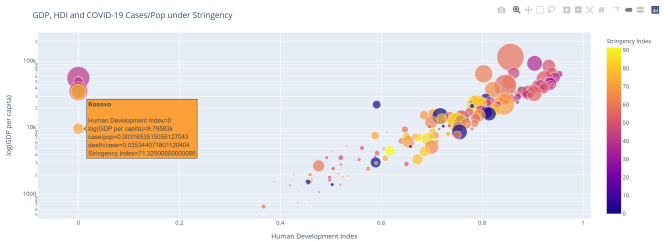

Fig. 2. Scatter plot showing 210 economic entities. Along the x-axis is the HDI and along the y-axis is logarithmically-scaled GDP per capita. The stringency index, indicated by the color bar to the right, represents the strictness of COVID-19 guidelines. The size of the scatter point represents the number of confirmed cases per population. The visualization is interactive, as shown above (example: Kosovo).

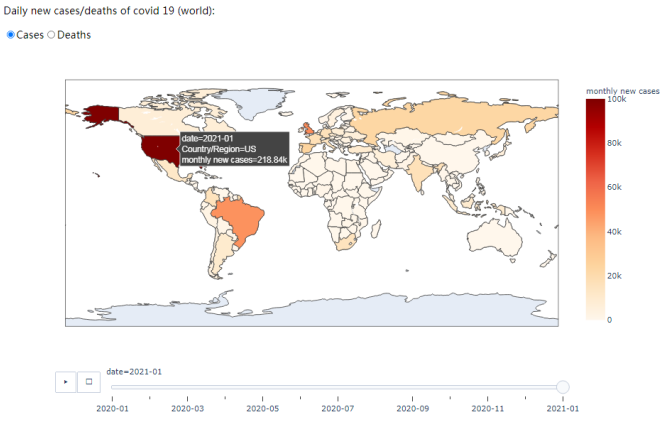

Fig. 3. Animated and interactive map plot showing the cases and deaths in the world from 2020-1 to 2021-1

lockdown policies a country, which has no obvious relationship with its economic status and epidemic situation from the graph. In order to confirm our deduction, we additionally explored global flight data to represent trade, traffic, and tourism conditions. Unfortunately, we did not find a suitable dataset covering the epidemic period, so we did not include it in the report. However, through our rough analysis, we found that many countries with low case/pop ratios didn’t even have flight records (which does not mean that there are no airports but that there are too few observations leading to missing data). Obviously, the traffic between developed economies is more convenient, and more frequent people flow, thus providing conditions for the spread of the virus among the population

3.2 U.S.

It can be seen from the global animation that the United States is one of the countries with the most severe COVID-19 pandemic. The new cases and deaths in the United States are rising at an alarming rate. From the animation of the U.S., the spread of COVID-19 shows a clear geographical trend, starting from densely populated urban centers and then spreading to rural areas. The epidemic is particularly severe on the economically developed east and west coasts, And the Los Angeles area has reached a peak in new cases and deaths in the past month.

4 Discussion

4.1 Analysis of main visualization plot

From the whole project, it is a better choice to choose the United States, which is rich in all kinds of data. From the main visualization chart made by plotly dash, we can analyze the impact of the US economy on COVID-19 from multiple angles. Integrating the US case animation and the CFR part of the main visualization, it can be seen that from the fourth quarter to the present, even if the new cases are still on the

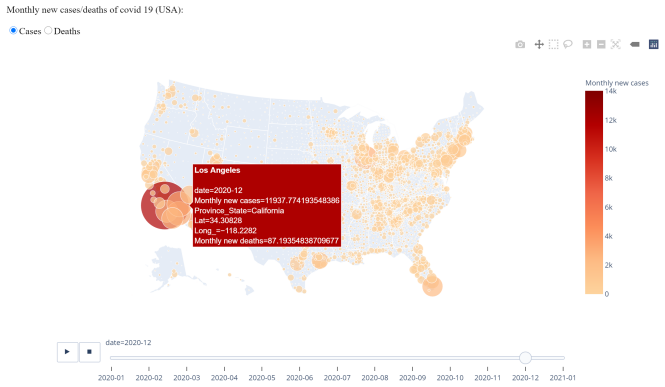

Fig. 4. Animated and interactive map plot showing the cases and deaths in the United States from 2020-1 to 2021-1

rise, the severity of Covid 19 has been declining, indicating that the Covid 19 epidemic is gradually being controlled in the United States and unemployment rate also shows a certain degree of geographic correlation with CFR.

Regarding personal income, we have observed a very interesting phenomenon, that is, in the second quarter when the epidemic became more severe, personal income rose greatly, and then fell sharply in the third quarter. After consulting the information, we speculate (https://www.bea.gov/news/blog/2020-05-29/personal-income-and-outlays-april-2020) this is due to government sent stimulus checks to households (2020 April), which are intended to stimulate the economy by providing consumers with some spending money. And this also reflects that the increase in government social welfare is sufficient to offset the decrease in most other sources of personal income. After that, the government's social welfare for individuals decreased and the epidemic raged, resulting in a substantial reduction in personal income changes. The possible reason for the substantial increase in GDP in the third quarter was that people's desire to consume was suppressed due to the epidemic and government assistance helped boost spending after disruptions caused by COVID-19 eased. Before the epidemic, US GDP growth had been steadily rising and the unemployment rate and PIC were also relatively stable. In terms of all aspects of the main visualization, the epidemic has greatly affected the policies and economic conditions of the United States.

4.2 Thoughts and expectation

When we initially began this project, we knew we wanted to look at the world's economic reaction to the COVID-19 pandemic. However, simply showing the number of cases/deaths and global GDP seemed barren and tedium. There obviously was much information missing since each country seemed to respond differently in their actions. We, therefore, decided to research other datasets to complement and supplement our primary dataset, the Impact of COVID-19 Pandemic on the Global Economy (Tiwari, 2020). An issue we faced when finding a primary dataset to build off of was that the chronological range of data was sometimes limited (e.g., data was only recorded until November rather than the entire year of 2020 for some datasets) and therefore did not satisfy our idealized goal. We eventually settled with the most abundant dataset we could find for the year 2020, the aforementioned dataset, as a baseline to build off of. We found governmental and academic-based databases to be most useful in our construction since they tend to be complete, well-maintained, and thorough. From there we were able to extrapolate visualizations to allow viewers to not only explore our directed interests but to view aspects of the pandemic that have not previously been articulated. Nevertheless, only limited data from Bureau of Economic Analysis (BEA) contain full data by each state, not to mention project the nationwide data on the U.S. map.

Based on these various datasets, our original plan was to use the choropleth map as backbone, projecting different data types (COVID-19, economy) on the map, and switching between different options by click interactions, so that users can observe and understand the differences and changes of data. This completion is relatively satisfactory. Obviously, it provides the functions we need - the micro-and macro-economic conditions under the epidemic, and the comparison between the status of individuals and the entity country are more completely reflected.

5 Conclusion

In this report, we selected the United States, where the Covid 19 epidemic is more serious and rich in data resources, as the main research object. Although there have been many visualization projects on covid19, we still hope to analyze from the aspects we are interested in. So we integrated economic-related data such as GDP growth, unemployment rate and personal income changes with CFR epidemic data, aiming to explore the impact of the US epidemic on the economy in a more integrated way.

However, the project still has limitations. As an improvement, we can add more data options, such as personal consumption, personal savings, GDP is only a comprehensive indicator, and does not refer to a certain aspect. For time selection, perhaps we can try range select, but quarterly data is different from monthly data (we tried by day, but the data set is too big and the frame collapsed), which is still to be considered. In addition, we can even add more interactions and views. For example, not limited to hovering, we can increase area selection and make statistics, and display the results in an additional graph, or we can add a line chart or a pie chart to separately display this part of the data, and describe the two types of data The relevance between them is set to the x and y axes, which are additionally displayed next to the map, and are not limited to the switching of single data.

6 References

Coronavirus Government Response Tracker. (2020, March). Retrieved January 18, 2021, from https://www.bsg.ox.ac.uk/research/research-projects/coronavirus-government-response-tracker.

Goldbloom, A. (2020, November 2). COVID-19 data from John Hopkins University. Retrieved January 18, 2021, from https://www.kaggle.com/antgoldbloom/covid19-data-from-john-hopkins-university/metadata.

Harrington, R. A. (2016, March 16). Britannica - Case fatality rate. Retrieved January 20, 2021, from https://www.britannica.com/science/case-fatality-rate

Human Development Reports. (n.d.). Retrieved January 18, 2021, from http://hdr.undp.org/en/content/human-development-index-hdi

International Monetary Fund. (2021). Retrieved January 19, 2021, from https://www.imf.org/external/datamapper/NGDP_RPCH@WEO/OEMDC/ADVEC/WI

Tiwari, S. (2020, November 29). Impact of Covid-19 Pandemic on the Global Economy. Retrieved January 19, 2021, from https://www.kaggle.com/shashwatwork/impact-of-covid19-pandemic-on-the-global-economy.

Tracking America's recovery. (2021, January 22). Retrieved January 22, 2021, from https://edition.cnn.com/business/us-economic-recovery-coronavirus

7 Appendix

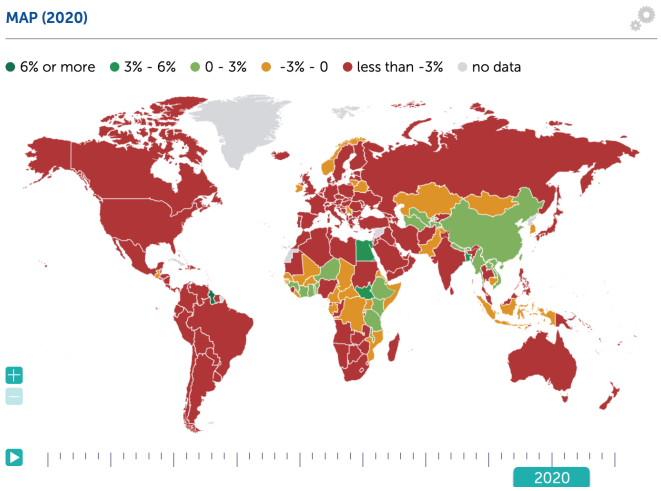

Fig. 5. Taken from the International Monetary Fund's website for real GDP growth, this figure shows the global economic impact with respect to GDP growth (in percentages) in 2020 owed to the COVID-19 pandemic (International Monetary Fund, 2021).

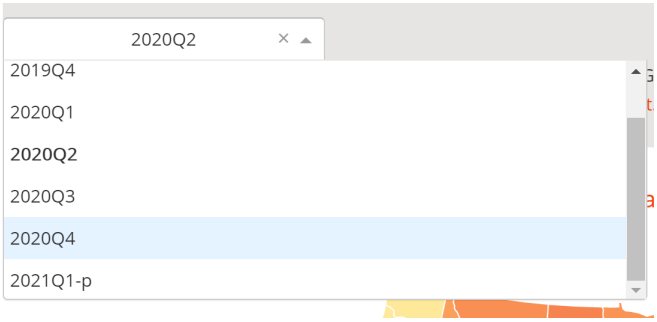

Fig. 6. A closer look at the drop-down menu for selecting fiscal quarters.

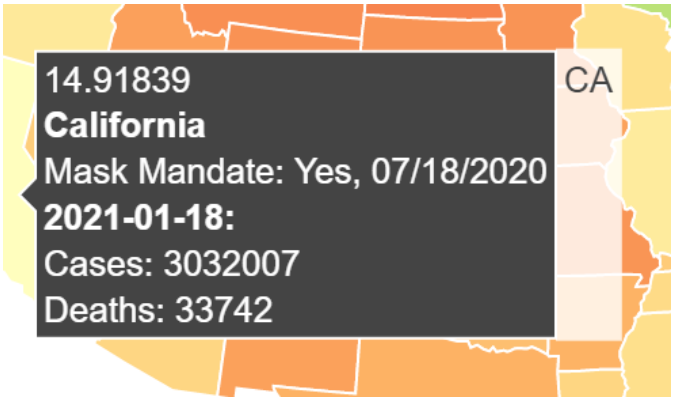

Fig. 7. A closer look at the interactivity of the dash-plot by hovering your cursor over a given state. Pictured here is California.

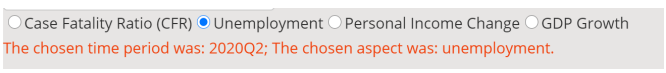

Fig. 8. selection of main visualization

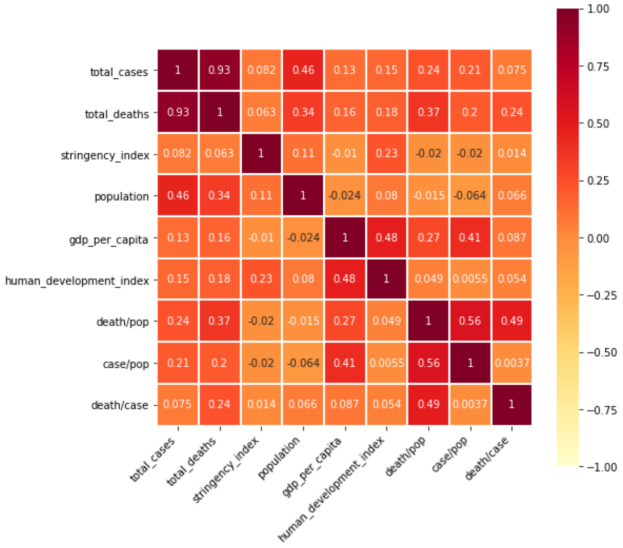

Fig. 9. correlation heatmap between COVID-19 cases or deaths and economic factors worldwide (mean values among all 210 countries or regions), HDI and case/pop are quite related to GDP per capita.