

Economic Impact of COVID-19 in the U.S.

Disha Srivastava and Rinija Raja
May 4th, 2021

Part 1: Introduction

Project Overview and Goal

Our project goal is to examine and understand the impact of COVID-19 on integral aspects of the U.S. economy such as GDP and Unemployment Rates; we chose these parameters as they are often used by economists to measure the current state of an economy.

We have first gathered COVID-19 health data recorded daily by the U.S. Center for Disease Control to understand which months of the pandemic faced the highest number of new cases and deaths. We utilized R to wrangle this data into monthly, national data and created bar plots to help us visualize the peaks and troughs in the pandemic timeline beginning with Jan '20 through Mar '21.

To study unemployment rates, we gathered Labor Force Statistics from the Current Population Survey Dataset from the U.S. Bureau of Labor Statistics. This dataset helped us understand the shifts in unemployment rates and the 1-monthly change in rates from Jan '20 to Mar '21. We mapped a timeline of the increase/decrease in cases and deaths along with the change in unemployment rates in order to infer a correlation. In creating these line charts, we calculated new cases and deaths as a proportional of the total (from Jan '20 to Mar '21) in order to place all three variables (unemployment rates, new cases, and deaths) on one axis.

In order to further understand economic impacts in specific regions of the U.S., we have focused our GDP analysis on four states with contrasting population sizes and responses to the pandemic. We specifically examined California (strong response to the pandemic, large population size), Texas (relaxed response to the pandemic, large population size), Rhode Island (strong response to the pandemic, small population size), and Utah (relaxed response to the pandemic, small population size). We also explored the GDP changes between various industries such as art, health, technical professions, real estate, and food to better understand which contributors to state GDP were impacted the most. In our analysis of GDP, we created an interactive map with data from 2019 and 2020 in order to compare pre and post pandemic GDP. We additionally created barplots that displayed the change in GDP from 2017 to 2020 for our four chosen states.

It has been saddening to study this data, quantifying the rise in cases and deaths and amount of individuals who lost their employment or seeing the downward slope of GDP, however, we hope our findings will help inform how government responses to the pandemic impacted local and state economies and the health of residents in their respective economy.

#Economy Visualization	Code
------------------------	------



The global pandemic this past year has impacted many aspects of our normalcy and daily lives. We utilized the Twitter Developer API to better understand how people across the world were thinking about the current state of our economy, by gathering the 600 most recent tweets published with #economy and creating a word cloud to visualize our findings. The top words in the word cloud are "covid", "business", followed by "jobs", "government", and "globalhealth" to name a few.

Hypothesis

We predict a linear correlation between unemployment rates and number of cases. As the number of COVID-19 cases increased, local and state governments entered lockdowns and further quarantine restrictions which impacted business and services, leading to higher levels of unemployment. In contrast, we hypothesize that an increasing number of deaths may not have had an equivalently strong correlation with unemployment rates. There are a few reasons for this: (1) cases spread and are recorded more quickly, (2) the number of cases is drastically higher than number of deaths, (3) there seemed to be a long-term misconception that only the elderly and immunocompromised (who might represent a smaller subset of the general population) were at high risk for severe loss due to the virus.

In terms of GDP per State, we have chosen California, Texas, Rhode Island, and Utah to examine in detail. We predict that California might have experienced an increase in GDP, due to the number of technology companies residing in the State of California who reported record amounts of sales during the shift to a remote, virtual world. We briefly researched this topic with the California State Assembly (<https://ajed.assembly.ca.gov/keyinsightsintoinnovation>). California and Texas each have large populations and were at the top of the charts at distinct points of time for cases/deaths, but followed drastically contrasting responses to the pandemic in terms of lockdowns, mask mandates, etc. We hypothesize that California would have seen an increase or at least continuation of GDP from previous years whereas Texas might have experienced a decline. Rhode Island and Utah are closely related in population amounts, however, Utah was the one of the only states to not issue a lockdown order. Specific counties in Utah with larger populations or tourist sites did issue lockdown orders, whereas Rhode Island's governor issued a stay at home order in late March itself. We briefly researched stay at home orders through the New York Times (<https://www.nytimes.com/interactive/2020/us/coronavirus-stay-at-home-order.html>). We predict that Rhode Island will experience a steady GDP in 2020 whereas Utah might experience a decline due to tourist locations and highly populated areas (with the highest job density) were the only ones experiencing lockdowns.

Variables

variable	description
U.S. COVID-19 Cases	Number of COVID-19 positive cases recorded each day in each state; this data has been wrangled into monthly national data from January 2020 to March 2021.

variable	description
U.S. COVID-19 Deaths	Number of deaths from COVID-19 recorded each day in each state; this data has been wrangled into monthly national data from January 2020 to March 2021.
U.S. Unemployment Rate per Month	National recorded unemployment rates on a monthly basis from January 2020 to March 2021.
U.S. GDP Percent Change per Year	National monthly change in unemployment rates on a monthly basis from January 2020 to March 2021.
Industry	Arts, Health, Food, Real Estate, and Technical.
States	California, Texas, Rhode Island, and Utah.

Part 2: COVID-19 Data Analysis

Visualization Code

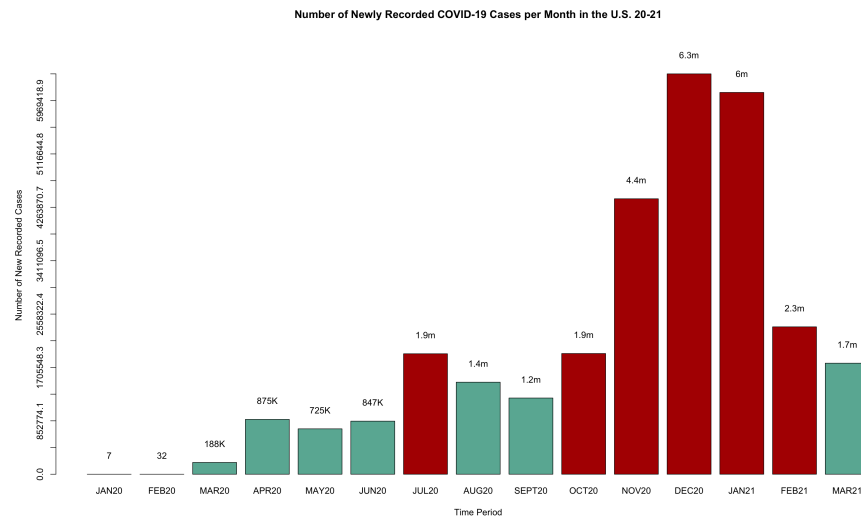
Data

Let's take a look at the total counts of new COVID-19 cases and deaths recorded each month in the U.S. This data has been reported by the U.S. Center for Disease Control in their data set: "United States COVID-19 Cases and Deaths by State over TimeCase Surveillance" (<https://data.cdc.gov/Case-Surveillance/United-States-COVID-19-Cases-and-Deaths-by-State-o/9mfq-cb36>). We have cleaned and transformed daily COVID-19 reported data by state into national data organized by month.

```
## months dates newcases deaths
## 1 JAN20 2020-01-01 7 0
## 2 FEB20 2020-02-01 32 0
## 3 MAR20 2020-03-01 188138 3680
## 4 APR20 2020-04-01 875947 55007
## 5 MAY20 2020-05-01 725178 42053
## 6 JUN20 2020-06-01 847821 21634
## 7 JUL20 2020-07-01 1924412 28249
## 8 AUG20 2020-08-01 1470219 29184
## 9 SEPT20 2020-09-01 1217378 22327
## 10 OCT20 2020-10-01 1928056 24012
## 11 NOV20 2020-11-01 4400344 39397
## 12 DEC20 2020-12-01 6395806 79189
## 13 JAN21 2020-01-01 6098794 97095
## 14 FEB21 2020-02-01 2354530 63431
## 15 MAR21 2020-03-01 1773722 33041
```

Now, let's look at how counts of new cases per month arose and fell from Jan '20 to Mar '21:

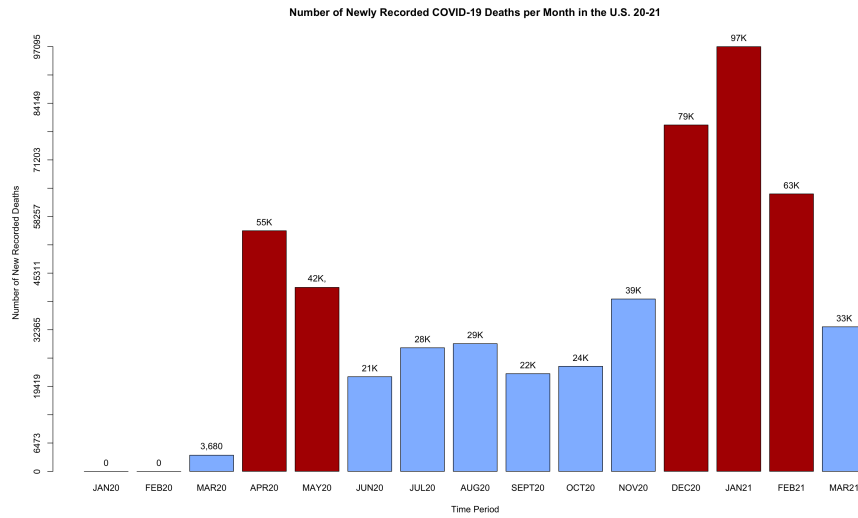
New COVID-19 Cases per Month



This bar plot helps us visualize a timeline for the spread of COVID-19. The red bars denote the top six months which experienced the highest number of recorded cases, often referred to as peaks or "second, third" waves. We can infer that July '20, Oct '20, Nov '20, Dec '20, Jan '21, and Feb '21 had the highest number of recorded cases; from this data, we might predict that unemployment rates would be higher during these months and GDP would be lower.

Now, let's look at how counts of new deaths due to COVID-19 per month arose and fell from Jan '20 to Mar '21:

New COVID-19 Deaths per Month



This bar plot reports the months which recorded the highest number of deaths due to COVID-19; the bars in red signify the top five months with the highest amount of deaths. These months include: Apr '20, May '20, Dec '20, Jan '21, and Feb '21. In correlation with the highest number of cases recorded, the months of Dec '20 through Feb '21 had both the highest number of cases and deaths. It is surprising that April and May of 2020 had two of the highest number of deaths although they didn't have a peak in cases. However, Apr/May were early months of the pandemic where we had little knowledge about the virus, and their average case count is around 800K which might have propelled the peak in July of 1.9 million cases.

Part 3: Unemployment Data Analysis

Unemployment Analysis and Visualizations Code

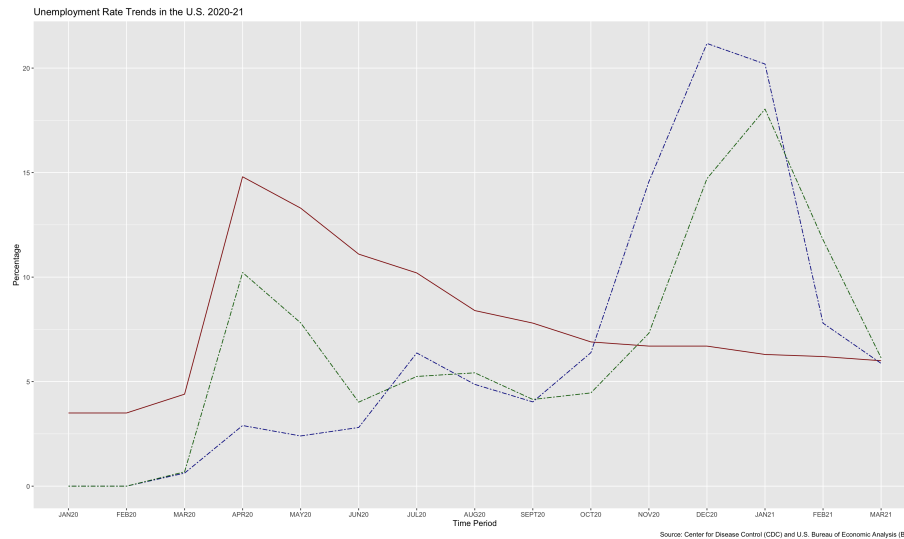
Data

The table below summarizes the unemployment rates and monthly change in unemployment per month from Jan '20 to Mar '21. New cases and deaths per month have been calculated as a proportion of the total cases and deaths from Jan '20 to Mar '21 in order to create a line chart with all three variables.

##	Date	Unemployment_Rate	Monthly_Change	percentcases	percentdeaths
## 1	JAN20	3.5	-2.8	0.00002317851	0.00000000
## 2	FEB20	3.5	0.0	0.00010595892	0.00000000
## 3	MAR20	4.4	25.7	0.62296558878	0.6836349
## 4	APR20	14.8	236.4	2.90044987507	10.2186703
## 5	MAY20	13.3	-10.1	2.40122112355	7.8122010
## 6	JUN20	11.1	-16.5	2.80731860893	4.0189560
## 7	JUL20	10.2	-8.1	6.37214414227	5.2478270
## 8	AUG20	8.4	-17.6	4.86821293398	5.4215222
## 9	SEPT20	7.8	-7.1	4.03100172501	4.1476949
## 10	OCT20	6.9	-11.5	6.38421021402	4.4607179
## 11	NOV20	6.7	-2.9	14.57049022953	7.3187949
## 12	DEC20	6.7	0.0	21.17789628105	14.7109692
## 13	JAN21	6.3	-6.0	20.19442534241	18.0373733
## 14	FEB21	6.2	-1.6	7.79635782115	11.7835998
## 15	MAR21	6.0	-3.2	5.87317697682	6.1380385

Now, let's take a look at correlations between unemployment rates, cases, and deaths:

Visualize Unemployment Rates' Timeline

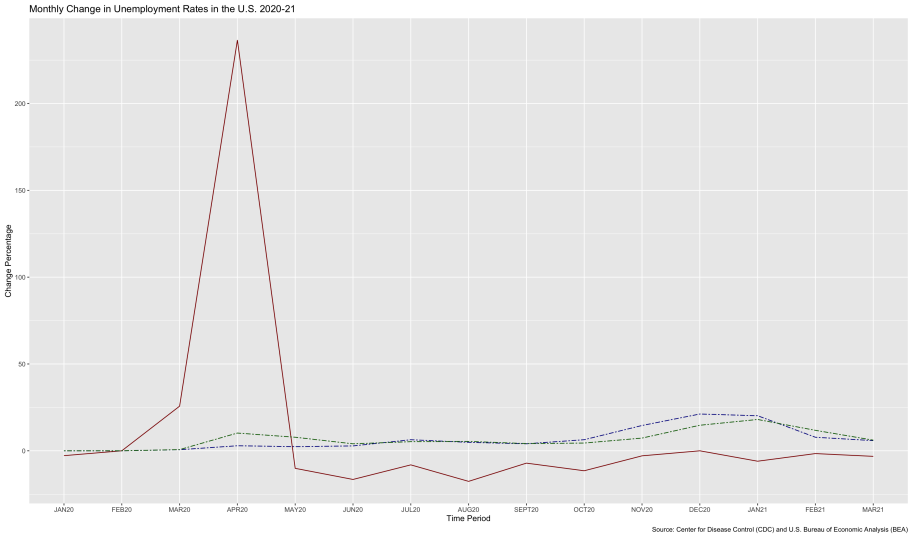


In this line chart, dark red represents **unemployment rate**, dark blue represents percent contribution of **new cases**, and finally, dark green represents percent contribution of **new deaths**.

This graph helps us visualize the correlation between unemployment rates and the spread of COVID-19 in terms of newly recorded cases and deaths each month from Jan '20 to Mar '21. In the beginning of the timeline, we see stagnancy in unemployment during Jan '20 through Feb '20, as the pandemic had not yet reached the United States (although we had 39 cases, it had not reached major headlines yet and people were mostly unaware of the gravity of the virus, and we were at 0 deaths). Mar '20 is when COVID-19 hit the news headlines, cases shot up, and we experienced our first brutal month of deaths from this pandemic. The graph visualizes this categorial phenomenon, as Apr '20 had the highest recorded unemployment rate of 14.8% due to lockdowns and quarantine orders affecting businesses nationwide. We have studied which industries were hit the hardest further down in our analysis. Apr '20 through Jul '20 we experienced a fluctuating increase and decrease in the number of new cases and deaths, with unemployment numbers climbing and falling in sync. The timeline from Sept '20 to Jan '21 is especially surprising, as the number of cases and deaths hit their highest peaks yet unemployment rates increased slightly and stayed mostly continuous on a downward slope. This may be due to most local and state governments easing lockdown restrictions and as a result most businesses reopening with a better understanding of how to operate safely (outdoor dining, sanitization routines, plexiglass dividers, etc.). Overall, we can infer from this line chart that in the early stages of the pandemic there was a correlation between number of cases and deaths and unemployment rate, as cases/deaths informed decisions by government officials for safety procedures such as lockdowns; however, in the late stages of the pandemic, we learned how to live with the virus and our business adapted new ways to keep customers safe in order to reopen after a devastating year of closures which lowered unemployment rates although cases/deaths were still climbing. This might also infer that we opened our economy too soon, as Oct '20 to Jan '21 is when lockdown restrictions eased and cases/deaths climbed drastically, but that is a debate for another day.

Now, let's take a look at correlations between monthly percent change in unemployment rates, cases, and deaths to confirm our graph from above:

Visualize Monthly Change Timeline



In this line chart as well, dark red represents **unemployment rate**, dark blue represents percent contribution of **new cases** , and finally, dark green represents percent contribution of new **deaths**.

From the timeline developed by this line chart we can infer three notable changes: first, the initial spike in unemployment from Feb '20 to Mar '20; second, the peak of unemployment change from Mar '20 to Apr '20; and finally, third, the period of constant fluctuation in changes to unemployment levels from May '20 to Mar '21. Looking at the monthly changes in unemployment helps us confirm our earlier findings: unemployment rates are directly correlated with new cases and deaths in the early stage of the pandemic, but are unstable in later stages. In Apr '20 almost every local and state government in the U.S. entered lockdowns, creating the initial spike in unemployment rates and the large monthly change from Mar '20 to Apr '20. In the period of May '20 through Mar '21, individual local and state governments have had varying responses to the pandemic in terms of lockdowns, mask mandates, and reopening guidelines, which makes it difficult to map a constant pattern in unemployment rates during this time period (which we might consider the later stages of the pandemic). In order to look in more detail at the economic impact of COVID-19 on specific regions with specific pandemic responses, let's now take a look at GDP by state and industry.

Part 4: GDP Data Analysis

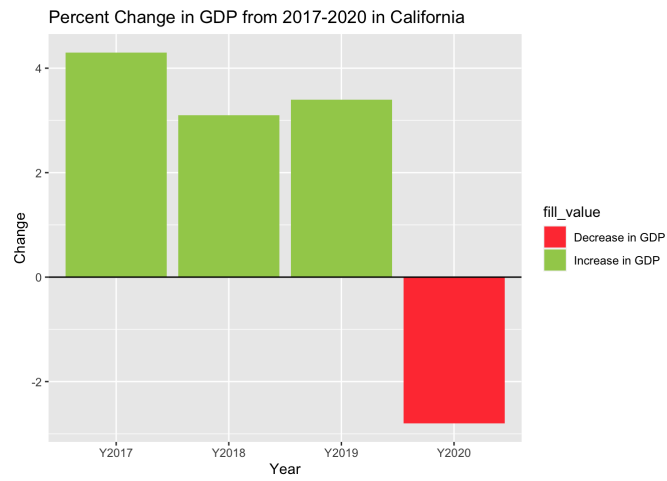
GDP Analysis and Visualizations	Code
---------------------------------	------

Data

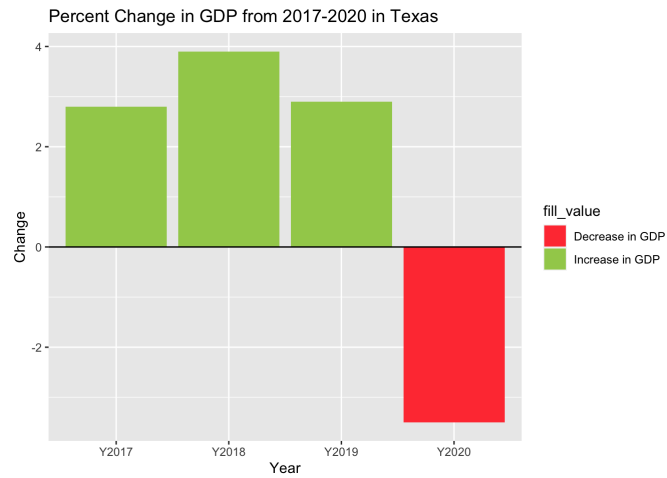
##	State_Name_Col	State_2020_Rank	Year_Col	GDP_Change_Col
## 1	California	15	Y2017	4.3
## 2	California	15	Y2018	3.1
## 3	California	15	Y2019	3.4
## 4	California	15	Y2020	-2.8
## 5	Texas	23	Y2017	2.8
## 6	Texas	23	Y2018	3.9
## 7	Texas	23	Y2019	2.9
## 8	Texas	23	Y2020	-3.5
## 9	RhodeIsland	38	Y2017	-0.6
## 10	RhodeIsland	38	Y2018	0.8
## 11	RhodeIsland	38	Y2019	1.0
## 12	RhodeIsland	38	Y2020	-4.5
## 13	Utah	1	Y2017	4.4
## 14	Utah	1	Y2018	5.6
## 15	Utah	1	Y2019	3.8
## 16	Utah	1	Y2020	-0.1

Explore Yearly Change in GDP by State

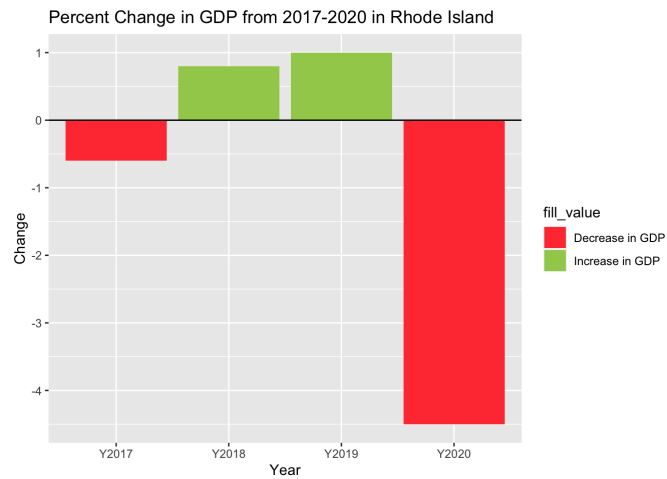
This is a set of four bar plots that display the percent change in GDP for California, Texas, Rhode Island, and Utah. Each graph includes data from one of the states mentioned above and displays the change in GDP from 2017 to 2020. Each bar in the chart has been filled with red or green to indicate whether there was an increase or decrease in GDP percentage.



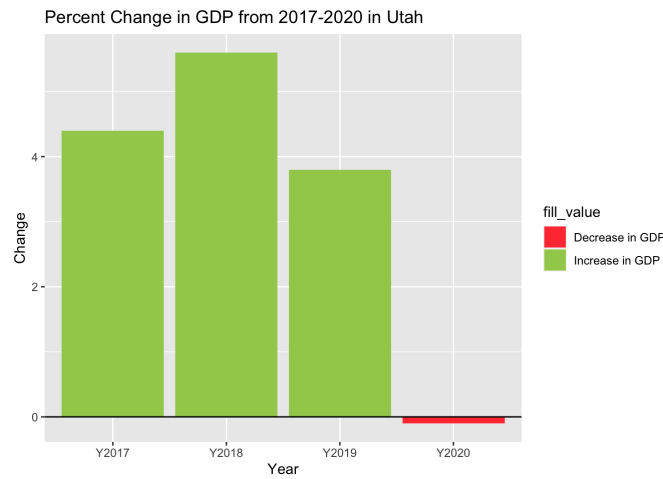
This bar plot displays the change in GDP for the state of California. We can conclude that the state had a very high starting percentage and the value dropped slightly in 2018 and 2019. However, the most drastic drop in GDP percentage happened in 2020 with a value of -2.3%.



This bar plot illustrates the GDP percent change in Texas. There is a starting percent change of roughly 2.5% and it significantly increased in 2018, but then dropped in 2019. Texas also had a significant decrease in 2020 with a value of -4.5%. This value is almost twice as low as California.



This graph displays the GDP change in Rhode Island from 2017 to 2020. This visualization is unique because there already was a GDP decline in 2017. Also, the GDP values for 2018 and 2019 were not as high as the other states. Lastly, Rhode Island had the highest drop in GDP for 2020 with a value of -4.5%.

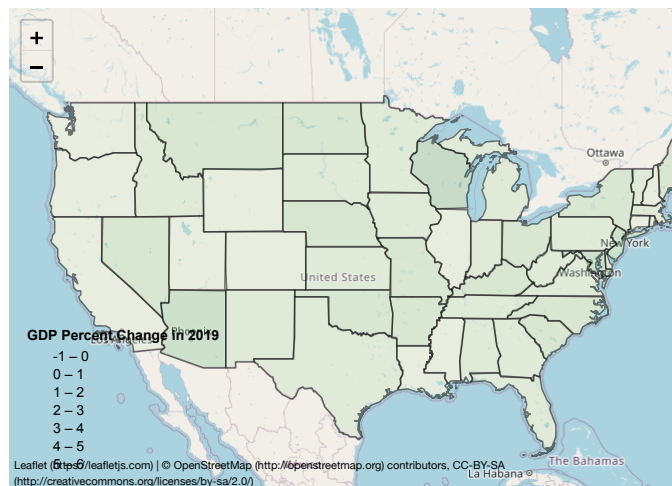


This last graph shows the change in GDP for the state of Utah. This state had the highest GDP values for 2017, 2018, and 2019. It is also important to note that there was a very small decrease in 2020.

Mapping Percent Change Visualizations

2019 2020

Explore the Percent Change in GDP by State in 2019

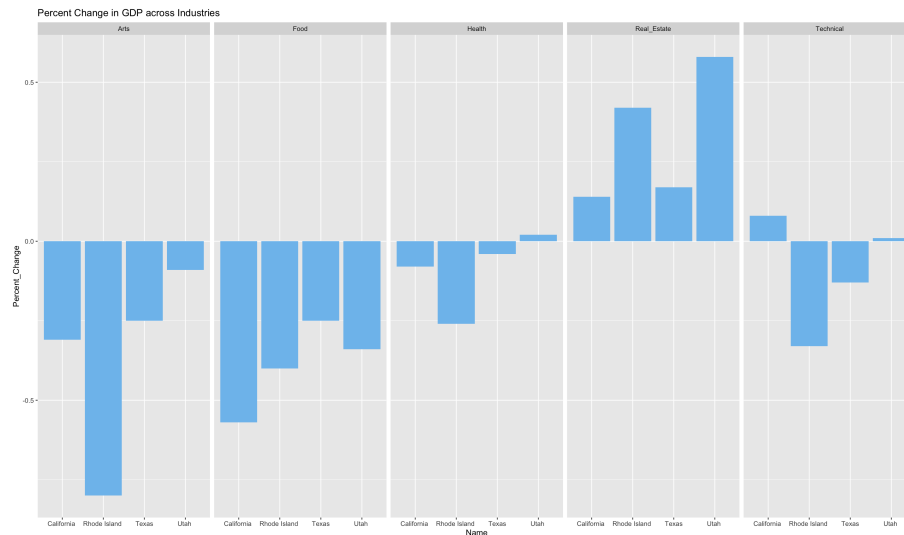


This is an interactive map of the United States where each state is colored in a green gradient based on the GDP percent change. We can conclude that the majority of states had a GDP value between 2 and 5 percent. It is interesting to note that there are no states with negative percentages because this was before the pandemic happened.

Visualize the GDP Change by State and Industry

Data

```
## # A tibble: 20 x 3
##   Name      Industry Percent_Change
##   <chr>      <chr>          <dbl>
## 1 "Rhode Island" Real_Estate    0.42
## 2 "Rhode Island" Technical    -0.33
## 3 "Rhode Island" Health      -0.26
## 4 "Rhode Island" Arts        -0.8
## 5 "Rhode Island" Food        -0.4
## 6 "Texas"      Real_Estate    0.17
## 7 "Texas"      Technical     -0.13
## 8 "Texas"      Health       -0.04
## 9 "Texas"      Arts         -0.25
## 10 "Texas"     Food         -0.25
## 11 "Utah"      Real_Estate    0.580
## 12 "Utah"     Technical     0.01
## 13 "Utah"     Health        0.02
## 14 "Utah"     Arts         -0.09
## 15 "Utah"     Food         -0.34
## 16 "California" Real_Estate    0.14
## 17 "California" Technical     0.08
## 18 "California" Health      -0.08
## 19 "California" Arts        -0.31
## 20 "California" Food        -0.570
```



This bar plot has been faceted to show the GDP percent change by each industry. The industries that had the lowest values were arts and food. The pandemic caused a lot of businesses to shut down and the arts industry was hit the hardest. It was interesting that California did not have a lower value because the majority of media/content is created in that state. The food industry was also significantly hit because many restaurants have closed or they have reduced their customer limits. It is interesting that Texas had a lower percent change because it is a largely population state. This could be attributed to the lower coronavirus restrictions resulting in less restaurants being closed. We can also observe that the real estate industry was doing significantly better than the other industries. Utah had the highest value so this means that a lot of people were buying properties in Utah when compared to the rest of the states. The technical industry was also interesting because California and Utah were the only states with an increase in value. It is surprising that the percentage change in GDP is very low because the technology hub of the country is located in California. During the pandemic, the major technology companies made impressive profits, so it is surprising as to why it is not reflected in this bar chart.

Part 5: Summary of Findings

GDP by State We were able to conclude that all states had a decline in their GDP during 2020. However, there were significant differences between the four states that we chose to examine. We found that Rhode Island had the largest GDP decline of -8% whereas Utah had a very minor difference of -1.7%. When we compared the heavily populated states, we found that California had a -2.3% decrease and Texas had a -4.5% decrease. It was surprising to see that there was such a low decrease in GDP for Utah in comparison to the other states, however, Utah was one of the only states in the U.S. which did not enforce a statewide lockdown. As we briefly studied earlier, in the resource published by the New York Times: "See Which States and Cities Have Told Residents to Stay at Home" (<https://www.nytimes.com/interactive/2020/us/coronavirus-stay-at-home-order.html>) only three counties in the state issued stay at home orders. Of those counties, Summit County is a popular tourist destination and Salt Lake City houses 1.2 million of Utah's 3.2 million residents. We can assume that most businesses and services were able to remain open throughout the pandemic in Utah, with the exception of tourist locations and their largest city, which is why they had the smallest decrease in GDP. California issued a stay at home order on Mar 19th whereas Texas waited until Apr 2nd, which might have caused the difference of almost a double in GDP declines. In terms of industries impacted by COVID-19, we were thoroughly surprised that real estate had soared during the pandemic in all states. All states saw an increase in arts and food, with only Utah seeing an increase in health services. We also noticed that California experienced an increase in the tech sector, which is most likely brought forth by Silicon Valley; California also experienced an increase in real estate, which might be due to tech workers looking to work from home full time and hence, looking for more spacious homes.

GDP Nationally When we analyzed the GDP changes for the United States, we were able to infer that state which experienced high amounts of growth in 2019 experienced high amounts of GDP loss in 2020. This may be due to construction projects and/or business and services which were pioneered in 2019 or booming in 2019 coming to a sharp halt in 2020 with unknown plans for reopening.

Unemployment Rate Timeline Unemployment Rate has been directly impacted by the COVID-19 pandemic, however, by creating line charts to help us visualize the timeline of unemployment rates, new cases, and new deaths, we were able to see a stark difference the early half of 2020 and the mid/latter of 2020 and early parts of 2021. In the beginning of the pandemic, most local and state governments had strong responses which enforced lockdowns and closures of businesses that were at high risk for spreading the disease (salons, preschools, restaurants, and more). In this time period the unemployment rate and monthly change in the first month of the pandemic (Mar-Apr '20) skyrocketed. As cases and deaths rose, more businesses shut down and more individuals were laid off (Airbnb, Tripadvisor layoffs). There was a compounding shift, however, towards the beginning of the summer of 2020 where unemployment rates began to consistently decrease regardless of cases and deaths increasing or decreasing. This may be due to businesses reopening under new CDC safety guidelines such as outdoor dining, plexiglass separations, the rise in demand for jobs such as delivery persons and other essential services. As we learned more about how to keep ourselves safe, more individuals were able to return to work or find new work, which decreases unemployment rates. In Oct '20 to Jan '21, the unemployment rate continues to cruise downward whereas cases and deaths reach an all time high. This may be due to state and local governments reopening too soon, providing individuals with unemployment again but with a higher risk of being infected with the virus (as cases and death rates were climbing during this period).

Conclusion Overall, this was an extremely interesting data science project to work on. We learned a tremendous amount about how to organize data into usable formats, especially how to work with different types of data such as time data, categorical reports, and connecting visualizations to reasons behind the trend lines. We hope our findings can help inform how the U.S. economy was impacted by COVID-19 specifically in terms of unemployment and GDP at a state and national level.

Part 6: Resources and Tutorials

United States COVID-19 Cases and Deaths by State over Time Dataset (<https://data.cdc.gov/Case-Surveillance/United-States-COVID-19-Cases-and-Deaths-by-State-o/9mfq-cb36>) from the U.S. Center for Disease Control and Prevention.

Labor Force Statistics from the Current Population Survey Dataset (<https://beta.bls.gov/dataViewer/view/timeseries/LNS14000000;jsessionId=C58F68E9140FE3B07E16AFFA7F77786E>) from the U.S. Bureau of Labor Statistics.

Gross Domestic Product by State, 4th Quarter 2020 and Annual 2020 (Preliminary) (<https://www.bea.gov/data/gdp/gdp-state>) from U.S. Bureau of Economic Analysis.

Interactive Choropleth Maps (https://learn.r-journalism.com/en/mapping/census_maps/census-maps/) from R Journalism.

R and Leaflet to create interactive choropleth maps (<https://towardsdatascience.com/r-and-leaflet-to-create-interactive-choropleth-maps-851ef83e275>) from Towardsdatascience.

Filter with Date data (<https://blog.exploratory.io/filter-with-date-function-ce8e84be680>) from Exploratory.io.

R Markdown: The Definitive Guide (<https://bookdown.org/yihui/rmarkdown>) from Bookdown.

Lesson 2. Twitter Data in R Using Rtweet: Analyze and Download Twitter Data (<https://www.earthdatascience.org/courses/earth-analytics/get-data-using-apis/use-twitter-api-r/>) from EarthLab.

How to Generate Word Clouds in R (<https://towardsdatascience.com/create-a-word-cloud-with-r-bde3e7422e8a>) from Towardsdatascience.

Part 7: Further Information

We have created a **GitHub Repository** (<https://github.com/dishsrivastava/economic-impact-of-covid19>) with our entire codebase and data sets for future reference.

Please note: we have also saved PNG images of our graphs and added them to the 'Graph-Images' folder in our GitHub Repository!