The Economic Impact of COVID-19

Chen An, Presnie Lu, Mingxuan Yang, Jishen Yin

Abstract

In this report, we analyze the economic factors that are influenced by COVID-19. In particular, we study several features, including the types of industry, and the change of GDP for each type at the state level. Finally, we build a model using PCA and the decision tree algorithm to do feature selection and predicting the economic growth rate in 1, 2, and 5 years for different cities.

Introduction

Since the global outbreak of the COVID-19 pandemic, scholars from research institutions and companies have begun to study the impact of the pandemic on the economy. From the aspect of the industry types, Dua, Ellingrud (2020) from McKinsey, and measured the vulnerability of US small businesses during the COVID-19 crisis concluded that accommodation & food service, health care, and construction are more vulnerable than other industries [1]. Tanaka, Higashide (2017) provided a bankruptcy model for forecasting the vulnerability of industrial economic activities [2].

This report was written from macro to micro, which started by comparing the impact between different countries. For the country which is mostly impacted by the COVID-19, we chose three representative states (CA, FL, NY) and focused on industry types and the trend of GDP change. The analysis of industry types is a novel point of our analysis. We also conducted specific analyses of key cities/ counties. Based on our analysis, we built a model that predicts economic growth in different regions using our selected features. Our model involves Principal Component Analysis (PCA), decision tree, feature selection, and data simulation.

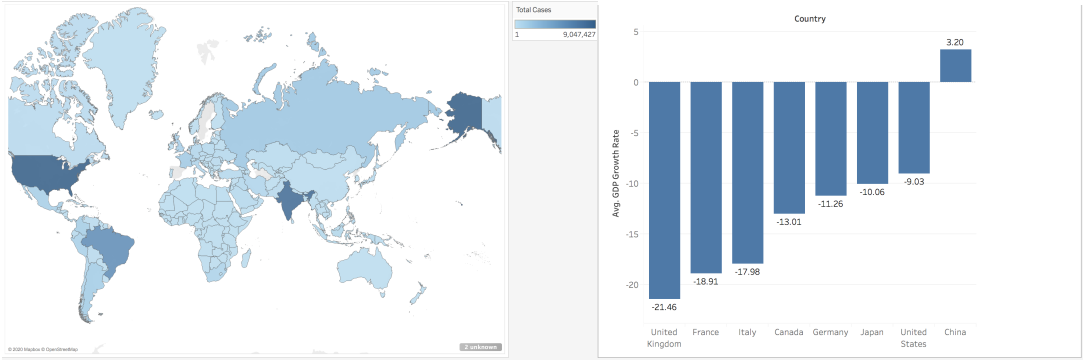
Methodology

This report is mainly about visualization. We used Tableau and pandas library from Python to build informative map plots, scatter plots, bar plots, pie charts, and prediction plots. Combined with conclusions from literature, we provided interpretations of these plots. We also used Principal Component Analysis (PCA), decision tree, feature selection, and data simulation to analyze the economic growth rate.

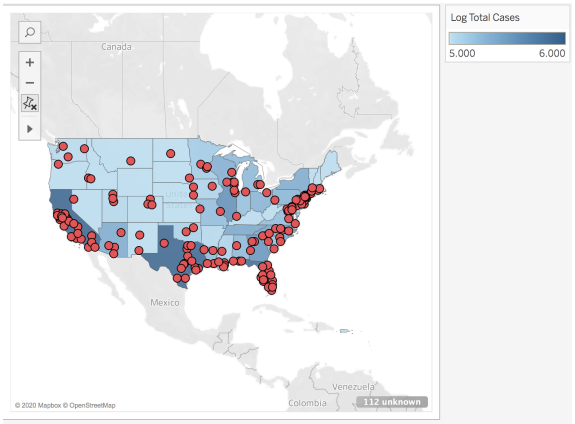
Analysis

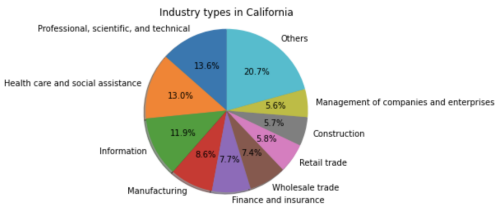
Our first graph (the left one below) shows the cumulative COVID-19 cases for each country. A country with a deeper color means more cases in that country. The data

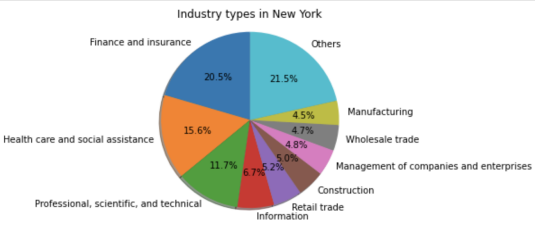
comes from [3]. It is obvious that the United States has the largest number of cases. To analyze the economic effect of COVID-19 around the world, we select eight major countries (G7 countries plus China) because the effect is the most obvious in countries with large economic size, Then we plot our second graph (the right one below) using the data source [6], which shows that the United States has the largest decrease of GDP in Quarter 2 2020 among all the eight countries. Because the United States has the largest number of cases, we concentrate on the states and cities in the United States for the following work.

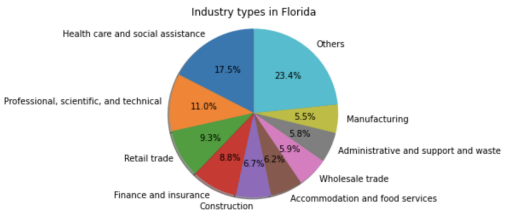


Our third graph (the upper left plot below) shows the cumulative COVID-19 cases at the state level, together with the first 300 counties with the most cases. The data comes from [4] and [5]. The sizes of the dots are proportional to the number of cases for each county. It is shown that the States of California, Florida, and New York have large numbers of cases. Thus, we choose the above three states as representatives to analyze the economic impact of COVID-19.



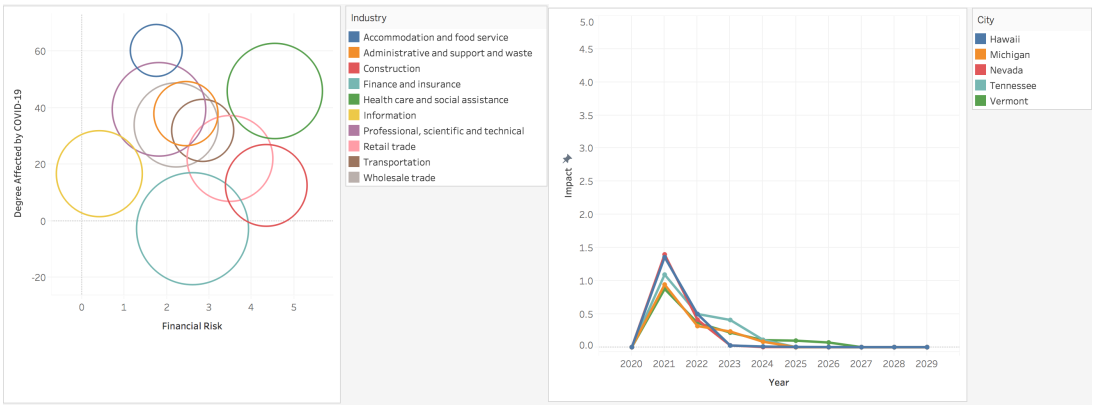






To understand the impact, we scrutinize the types of industries for each representative state. We obtain and plot pie charts (three charts above) of the key industries in these three states using data from [12]. Then we get the quarterly changes on the industry types from [8]. Notably, the quarter includes Q2 2020, where the covid began to have a deep effect on the economy. The key industries include Health care and social

assistance, Accommodation and food services, Administrative and support and waste, Manufacturing, Retail trade, and Construction. We use the data sources [8] and [10] to plot the graph on the types of industry about their financial risks and the degree affected by COVID-19 (the left one below). A larger size of a circle means that the corresponding type of industry has a larger share of GDP.



Finally, we build a model predicting the economic growth rate for a few regions. We first combine the factors of quarterly GDP (data source: [13]) and unemployment rate (data source: [9]), using PCA to obtain a factor that measures economic growth. With this PCA factor, we are able to select top cities that suffer from COVID-19 most and top cities that have the largest economic growth. Next, we use the decision tree algorithm on the variables of Industry (the combined contribution to GDP from the six key industries), population density (data source: [11]), total COVID-19 cases, and total COVID-19 related deaths, to obtain feature importance. From the model together with feature importance, we simulate and predict the economic growth rate for the top cities that suffer from COVID-19 most and top cities that have the largest economic growth (see the graph on the right above). As we shall see below, the five states correspond to those areas that are most negatively affected by COVID-19. Here I only provide a screenshot of python output as an indication. A lower PCA score identifies a more severe impact.

	State	industry	population2020	total unemp	unemp rate	population2019	Density	Total Cases	Total Deaths	Q1 2019	Q2 2019	Q3 2019	Q4 2019	Q1 2020	Q2 2020	PCA
27	Nevada	0.5997	2471346.0	227426.0	15.2	2421890.0	28.5993	98554	1769	0.5	1.8	4.7	2.7	-4.9	-42.2	-13.990139
10	Hawaii	0.5884	1089320.0	84481.0	13.4	1091223.0	219.9424	14841	214	0.1	-1.0	1.5	2.7	-8.9	-42.2	-13.573755
41	Tennessee	0.5499	5433837.0	305373.0	9.6	5375873.0	167.2749	256880	3263	1.1	1.8	2.5	-0.1	-3.9	-40.4	-10.943545
21	Michigan	0.5442	8008545.0	738476.0	14.9	7990795.0	177.6650	190043	7653	0.9	0.3	3.4	0.9	-7.9	-37.6	-9.445509
44	Vermont	0.6109	518890.0	32452.0	9.5	518151.0	68.1416	2155	58	3.8	-2.8	1.5	0.9	-5.8	-38.2	-8.780084

Findings

We have multiple observations from our model.

First, through the PCA, we know that a low unemployment rate contributes to 89% of the economic growth while a higher quarterly GDP contributes to 11% of the economic growth.

Second, we obtain the answers to which cities are affected by COVID-19 most. Although we do our modeling at the state or even country level, there is no loss of generality because we can pick the most representative city in the region. Our model

shows that the city that is most negatively affected by COVID-19 is Las Vegas. This is plausible because Las Vegas is famous for its gambling industry and tourism. Due to COVID-19, casinos are either closed or restricted and much fewer people would go gambling or taking a trip to Las Vegas. Our model also shows that the city that grows fastest due to COVID 19 is Beijing. This matches our reality perfectly as China is the only major country that has an increase in the second-quarter GDP in 2020.

Third, in terms of feature importance, our finding is shown below as an output of python code. The number of deaths ranks the first, which means that the economic growth rate is mostly related to the fatality situation; the number of COVID-19 cases is the second; the industry type is the third; population density is the fourth. This finding makes sense because the number of deaths and the number of cases attract the most attention from people. The importance of industry type has already been shown in the example of Las Vegas. The population density is fairly vital as well, contributing to almost 20% of feature importance. For instance, NYC is one of the places that was struck by COVID-19 severely and it is probably related to its dense population.

```
Feature Importance of industry 0.1985
Feature Importance of Density 0.141
Feature Importance of Total Cases 0.2525
Feature Importance of Total Deaths 0.408
```

Fourth, from the second finding above, we know that the top five cities that feel the economic impact of COVID-19 most acutely are Las Vegas, Honolulu, Nashville, Detroit, and Burlington. For the predictions on these cities starting from 2021, the economic impact of COVID-19 will be the most severe in 2021 but will be mitigated in 2022, and there will be little impact in 2025.

Conclusion

As we are exploring the data, we find that the world is suffering from COVID-19 and it will take a long time to recover. It is worth mentioning that there is an increasing trend in the number of cases in the United States and worldwide in recent days. Still, based on our prediction plot, we are confident that the world will go back to the old normal after two or three years. Even for the cities that are affected most severely, they will recover in less than half-decade. Our analysis provides a few warnings, though. Based on our computation of feature importance, we realize that not only do the number of deaths and the number of COVID-19 cases play a role in the economic effect, but the type of industry also is not negligible. Besides, population density plays a role. Therefore, it is suggested that governments, no matter what size is, should make a diverse combination of industry types, so that the region is more likely to withstand the impact of COVID-19. In particular, the industry type of information is less affected by COVID-19 and has less financial risk, and the government should consider expanding the industry of information.

Appendix

Our work is uploaded to the following GitHub page:
https://github.com/QuietMorning/Duke_Datathon_2020

Reference & Data Source

[1] Dua, Ellingrud (2020). Which small businesses are most vulnerable to COVID-19 and when. McKinsey & Company
[2] Tanaka, Higashide (2017). Forecasting the Vulnerability of Industrial Economic Activities: Predicting the Bankruptcy of Companies. Journal of Management Information and Decision Sciences, 1532-5806
[3] https://ourworldindata.org/coronavirus-data
[4] https://covid.cdc.gov/covid-data-tracker/#cases_casesper100k
[5] https://usafacts.org/visualizations/coronavirus-covid-19-spread-map/
[6] https://stats.oecd.org/index.aspx?queryid=350
[7] https://github.com/nychealth/coronavirus-data
[8] https://fred.stlouisfed.org/release/tables?rid=331&eid=2088#snid=2122
[9] https://www.bls.gov/web/laus.supp.toc.htm
[10] https://csimarket.com/Industry/industry_Financial_Strength_Ratios.php?ind=903
[11] https://www.governing.com/gov-data/population-density-land-area-cities-map .html
[12] https://www.census.gov/data/datasets/2018/econ/cbp/2018-cbp.html
[13] https://www.bea.gov/news/2020/gross-domestic-product-state-2nd-quarter-2020