WPI

CS 573 Process Book

Analysing the Effect of U.S. State Government Policy Actions on the Spread of COVID-19

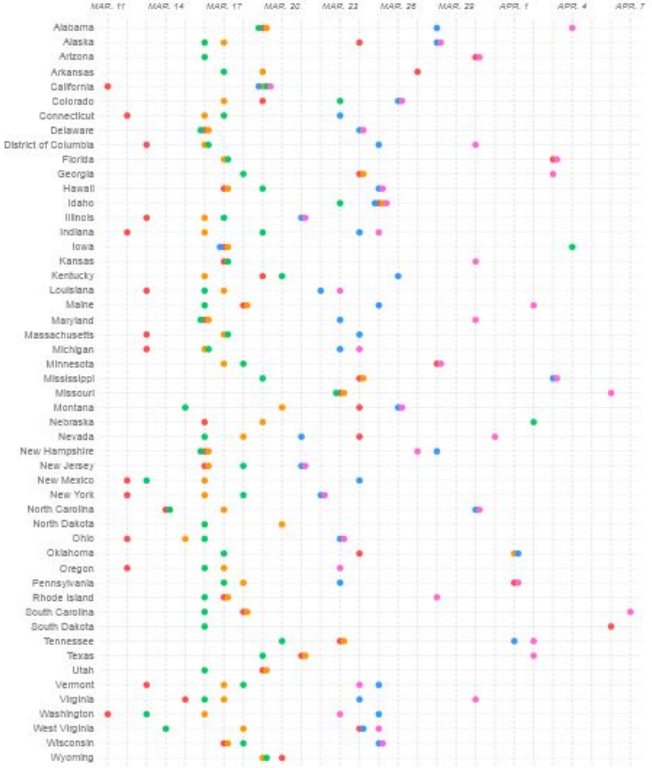
Narges Ahani

Amorn Chokchaisiripakdee

Effect of Social Distancing Policies Enactment Time on Spread of COVID-19

• Mass gathering restrictions • Initial business closure • Educational facilities closed • Non-essential services closed • Stay at home order

Policy Enactment Time



Ratio of Increase, One Week After Enactment Time

Overview & Motivation

In late 2019, Coronavirus and the disease caused by it, COVID-19 started to grow rapidly around the world. World Health Organization declared it a *pandemic* on March 11th, 2020 and many countries including United States declared the State of Emergency afterward.

Since then, many data analytics and data visualization techniques have been developed to keep track of data related to pandemic. These developments help people and governments to monitor the spread and effects of the virus and also understand the underlying story and make better decisions for near future and policies that should be enacted.

While many people visualized the direct data and trends regarding the actual number of people who have been affected by the virus worldwide, many works explore the indirect data including the impact of COVID-19 on different subjects in the world such as world economy or community mobility.

Moreover, some study conducted on analysing and visualizing the effect of different factors on the spread of COVID-19, such as life style of people in different regions or the time and severity of the preventive and controlling policies that was placed.

We also decided to focus on visualization of the data related to COVID-19 aimed at covering one aspect associated with the pandemic which will be discussed later in this report.

Related Work

As the focus of this project was to address the indirect data of COVID-19, we list the following works that inspired us to gain some insight and idea.

1. In Social Distancing Scorecard [1], author grades each state and county social distancing efforts by comparing mobility before and after COVID-19. Updated daily, this dashboard gives insight into how effective our communities have been in reducing travel and, in turn, slowing the spread of the virus.

2. State Data and Policy Actions to Address Coronavirus [2], not only visualized the number of confirmed, tested and hospitalized cases in different states in U.S. but also visualized the current condition (active:Yes/No) of different action policies in different states in U.S. They also published detailed tables of policies enactment times for each states. This work was the main inspiration for our project.

3. COVID-19 Mobility Data Network [3] uses aggregated and privacy preserving data from Facebook to display county level changes in the rates of mobility and staying put. This visualization provides guidance on how distancing efforts are being followed.

4. IHME COVID-19[4] projections, project the enactment time of different physical distancing policies ranging from school and restaurant closures to stay-at-home orders and the rates of change in mobility. It also visualize the trend of some related measures such as infections and tests and hospital resources use.

[1] Social Distancing Scoreboard
[2] State Data and Policy Actions to Address Coronavirus
[3] COVID-19 Mobility Data Network
[4] IHME COVID-19 Projections

Questions

In the absence of a vaccine or an effective treatment, controlling and slowing down the spread of the virus requires aggressive action from states and the federal government to reduce the spread of the virus and avoid overburdening on medical systems and ensure access to testing for those who need it and treatment for those who contract the disease.

States have taken a number of actions at different points of time aimed at reducing rates of mobility and avoiding large gatherings to mitigate the spread of the virus and reduce the likelihood that the virus is transmitted from person to person.

In this project we use visualization techniques to show the effect of policy enactment time on the spread of disease in each state. In particular we want to compare different states in U.S. and explore the variance in the number of confirmed cases, before and after the policies enactment time. This may help us to learn which policy is working better during the pandemic and if the earlier time of enactment positively help slow down the spread of the virus.

The policies that we studied in this report include mandatory stay at home orders, closures of non-essential businesses, bans on large gatherings, school closures, and limits on bars and restaurants and other public places.

The initial idea of analysing the variance in measures given mid-size government responses to COVID, was first proposed by Professor Lane Harrison. The initial driving question was that weather the variance in rates/deaths,etc. in US states different than provinces/states-equivalent in other countries? Due to limitation in gathering data, we modified the initial idea to analysing the variance across different states in U.S.

Questions

So the goal is to calculate and visualize the variance in the number of confirmed cases after enacting policies. This measure, alongside with the demonstration of the enactment time of each policy in different states, can facilitate the comparisons across states and help us understand whether the enactment time (earlier or later) effected on controlling the spread of the virus in different states.

Naturally, many factors influence on the variance including the lifestyle of people in each states, the absolute population of state, density (population per area) and crowdedness of the counties and regions in states and also having large transportation hubs (e.g. NY). However data shows that some states acted early and seem to have reduced spread (e.g. CA).

We didn't use the time-varying trend data. The reason is that even though these policies enacted, all states observed increase in the number of confirmed cases. So just looking at the trend data for different states is not informative enough for us to do the comparison and find the answer of our question. What we particularly are intended to know is *"when exactly a policy was enacted"* and *"how much increase were observed after that"*.

Thus we come up with a specific measure of variance described and formulated in table 1. In short this measure is the ratio of increase in number of confirmed cases, one week after enactment time of a policy in a state.

This measure takes into account the difference in population in states, as it is the ratio of increase in number of confirmed cases, after a policy was enacted.

Questions

Table 1: Notation for Measure of Variance Formulation

Parameter	Definition
$\mathcal{P}$	Set of enacted policies
$\mathcal{S}$	Set of states in U.S.
t	A period of time
$A_p^s(t)$	Number of confirmed cases after time t of enacting policy $p \in \mathcal{P}$ in state $s \in \mathcal{S}$
B_p^s	Number of confirmed cases at the time of enacting policy $p \in \mathcal{P}$ in state $s \in \mathcal{S}$
$\mathcal{V}_p^s(t)$	Measure of variance after time t for policy $p \in \mathcal{P}$ in state $s \in \mathcal{S}$

$$\mathcal{V}_p^s(t) = \frac{A_p^s(t) - B_p^s}{B_p^s}$$

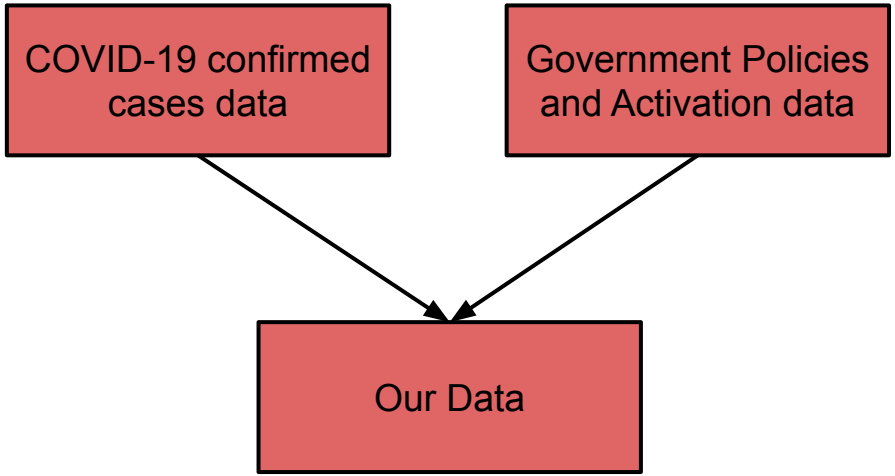
Data

We need to find 2 important data which are

1. General COVID-19 data which contains dates, states, confirmed cases, deaths etc.

2. Government Policies and activation dates

The first data is more general than the government policies data because there are many visualization working on the confirmed cases and death

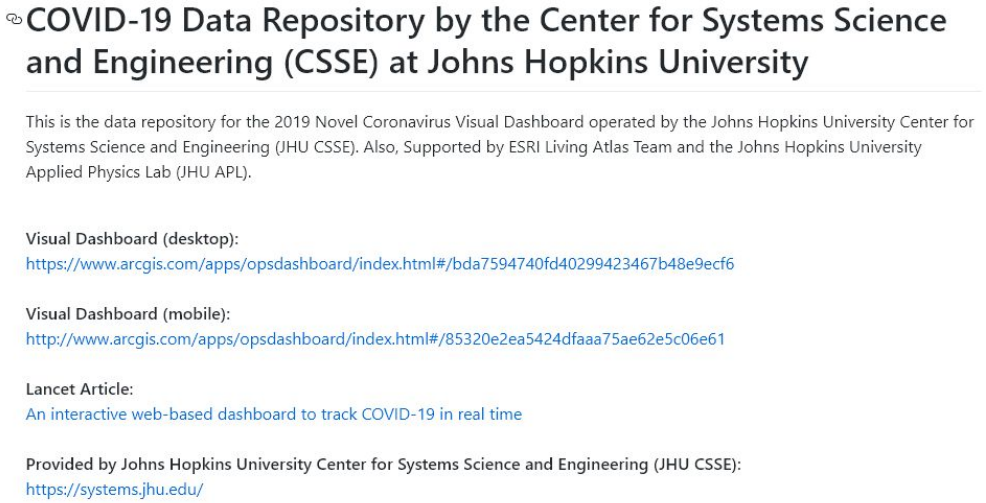


Data

We found the COVID-19 confirmed cases data from Github repository for the 2019 Novel Coronavirus Visual Dashboard operated by the Johns Hopkins University Center for Systems Science and Engineering (JHU CSSE)[1].

The particular dataset we used is named “time_series_covid19_confirmed_US.csv”

This dataset consists of states, county-level, longitude, latitude, and the number of confirmed cases from Feb 1st to May 1st.



COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University

This is the data repository for the 2019 Novel Coronavirus Visual Dashboard operated by the Johns Hopkins University Center for Systems Science and Engineering (JHU CSSE). Also, Supported by ESRI Living Atlas Team and the Johns Hopkins University Applied Physics Lab (JHU APL).

Visual Dashboard (desktop):
https://www.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e9ecf6

Visual Dashboard (mobile):
http://www.arcgis.com/apps/opsdashboard/index.html#/85320e2ea5424dfaaa75ae62e5c06e61

Lancet Article:
An interactive web-based dashboard to track COVID-19 in real time

Provided by Johns Hopkins University Center for Systems Science and Engineering (JHU CSSE):
https://systems.jhu.edu/

[1] https://github.com/CSSEGISandData/COVID-19

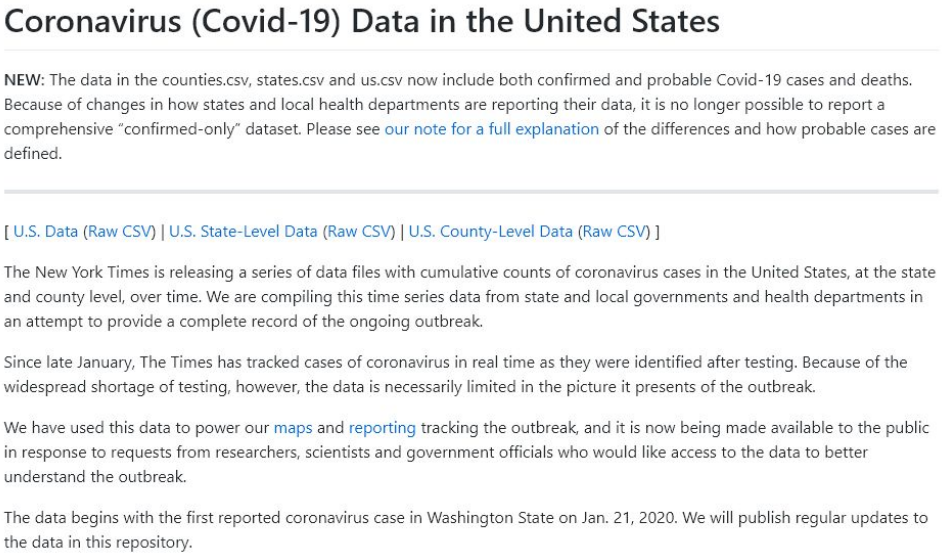
Data

New York Times Coronavirus (COVID-19) data in the United States

We found this dataset on github which contains 3 datasets :

1. U.S. Data
2. U.S. State-level Data
3. U.S. County-level Data

After reviewing both dataset from Johns Hopkins University and New York Times, we decided to use JHU dataset because it is more organized timeline than NYT.



Coronavirus (Covid-19) Data in the United States

NEW: The data in the counties.csv, states.csv and us.csv now include both confirmed and probable Covid-19 cases and deaths. Because of changes in how states and local health departments are reporting their data, it is no longer possible to report a comprehensive "confirmed-only" dataset. Please see our note for a full explanation of the differences and how probable cases are defined.

[U.S. Data (Raw CSV) | U.S. State-Level Data (Raw CSV) | U.S. County-Level Data (Raw CSV)]

The New York Times is releasing a series of data files with cumulative counts of coronavirus cases in the United States, at the state and county level, over time. We are compiling this time series data from state and local governments and health departments in an attempt to provide a complete record of the ongoing outbreak.

Since late January, The Times has tracked cases of coronavirus in real time as they were identified after testing. Because of the widespread shortage of testing, however, the data is necessarily limited in the picture it presents of the outbreak.

We have used this data to power our maps and reporting tracking the outbreak, and it is now being made available to the public in response to requests from researchers, scientists and government officials who would like access to the data to better understand the outbreak.

The data begins with the first reported coronavirus case in Washington State on Jan. 21, 2020. We will publish regular updates to the data in this repository.

Data

We first found the KFF[1] website which works on the state data and policy actions to address coronavirus.

This website contains some visualizations such as U.S. map confirmed cases & deaths by state. But the most important part is there are data of government policies.

KFF has two sections of policies :

1. State Social Distancing Actions
2. State COVID-19 Health Policy Actions

State Social Distancing Actions

View Full Table | Export CSV

Location	State Is Easing Social Distancing Measures	Stay at Home Order	Mandatory Quarantine for Travelers	Non-Essential Business Closures	Large Gatherings Ban	Sc
United States	Yes (34); No (17)	Original Stay at Home Order in Place (29); Stay at Home Order Eased or Lifted (16); No Action (6)	Original Traveler Quarantine Mandate in Place (20); Traveler Quarantine Mandate Eased or Lifted (3); No Action (28)	Original Non-Essential Business Closures Still in Place (17); Some or All Non-Essential Businesses Permitted to Reopen (28); No Action (6)	Original Gathering Ban/Limit in Place (40); Gathering Ban/Limit Eased or Lifted (9); No Action (2)	Closed (36); Re Recom School
Alabama	Yes	Lifted	-	Some Non-Essential Businesses Permitted to Reopen with Reduced Capacity	>10 People Prohibited	Closed
Alaska	Yes	Lifted	All Travelers	All Non-Essential Businesses Permitted to Reopen with Reduced Capacity	Expanded to 50+ People Prohibited	Closed
Arizona	Yes	Statewide	From Certain States	Some Non-Essential Businesses Permitted to Reopen with Reduced Capacity	>10 People Prohibited	Closed
Arkansas	-	-	-	-	>10 People Prohibited	Closed
California	Yes	Statewide	-	Some Non-Essential Businesses Permitted to Reopen with Reduced Capacity	All Gatherings Prohibited	Recom Year
Colorado	Yes	Rolled Back to High Risk Groups	-	Some Non-Essential Businesses Permitted to Reopen with Reduced Capacity	Expanded to >10 People Prohibited	Closed
Connecticut	-	Statewide	-	All Non-Essential Businesses Closed	Other [1]	Closed
Delaware	Yes	Statewide	All Travelers	Some Non-Essential Businesses Permitted to Reopen with Reduced Capacity	>10 People Prohibited	Closed
District of Columbia	-	Statewide	-	All Non-Essential Businesses Closed	>10 People Prohibited	Closed
Florida	Yes	Lifted	From Certain States	Some Non-Essential Businesses Permitted to Reopen with Reduced Capacity [2]	Other [3]	Recom Year
Georgia	Yes	Rolled Back to High Risk Groups	-	Some Non-Essential Businesses Permitted to Reopen with Reduced Capacity	>10 People Prohibited	Closed

[1] KFF,State Data and Policy Actions to Address Coronavirus

Data

The state social distancing actions is suitable to our project. However, the data they provide in csv format does not include the date of the policies.

We check on the sources of the data which has the data in plain text. We extract the date and policy to the text file.

Unfortunately, the policies we extracted from KFF are not useful to us due to the various policies in the data. And we do not have the measure to decide which policies to use in our project.

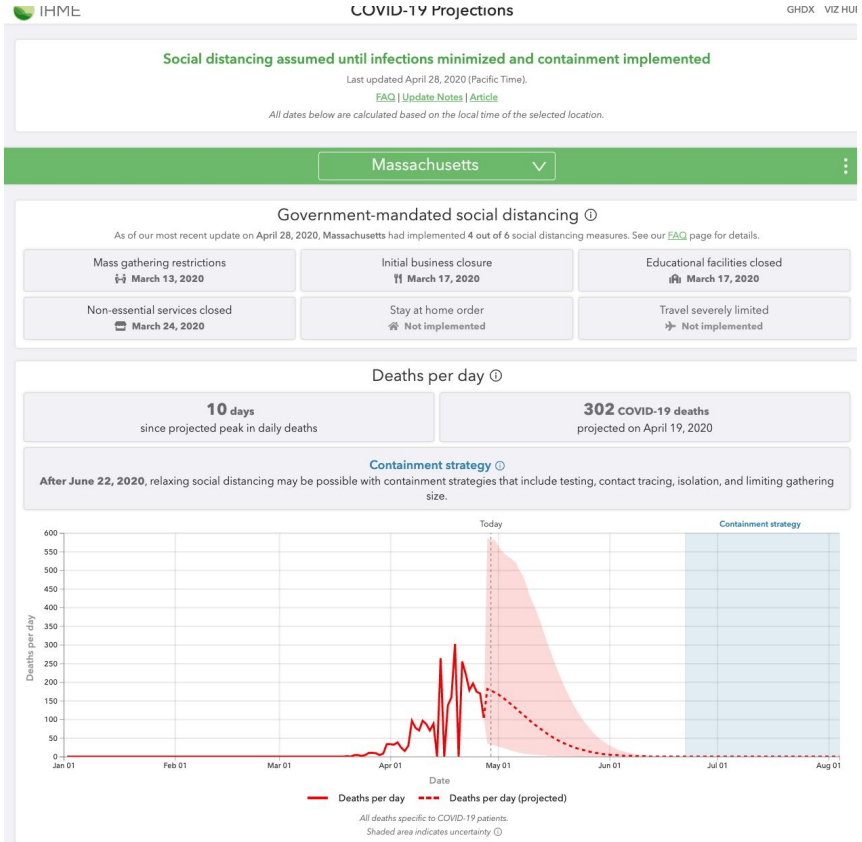
```
ALABAMA
3/13 Public Health Emergency
3/17, 3/19, 3/20, 3/26, 3/27 School Closures, Large Gatherings Ban, Bar/Restaurants Limits
3/18 Primary Election Postponement
3/23 Section 1135 Waiver
3/27, 4/3 Non-Essential Business Closures, Stay At Home Order

ALASKA
3/3, 4/15 Early Rx Refill
3/11 Public Health Emergency
3/13, 3/20, 4/9 School Closures
3/17, 3/23, 3/27 Stay At Home Order, Mandatory Quarantine for Travelers
3/17 Bar/Restaurant Limits
3/18, 4/15 Premium Grace Period
3/20, 3/23, 3/27 Non-Essential Business Closures, Large Gatherings Ban
4/2 Section 1135 Waiver
4/24 Easing Social Distancing Measures

ARIZONA
3/11 Emergency Declaration / Public Health Emergency
3/16, 3/20, 3/30 School Closures
3/19 Bar/Restaurant Limits
3/23 Section 1135 Waiver
3/30 Stay at Home Order
3/30 Large Gatherings Ban
4/7 Mandatory Quarantine

ARKANSAS
3/11 Public Health Emergency
3/15, 3/20, 4/4 School Closure, Bar/Restaurant Limits
3/20, 3/27 Premium Grace Period
3/26, 4/4 Large Gatherings Ban
4/2 Section 1135 Waiver

CALIFORNIA
3/11 Emergency Declaration
3/20 Marketplace SEP
3/12 Executive Order mandating compliance with state guidance
3/16 Large Gatherings Ban
3/16 Bar/Restaurant Limits
3/18 Early Rx refill
3/19 Stay At Home Order, Non-Essential Business Closures
3/23 Section 1135 Waiver
4/1 School Closures
```

Data

From the KFF data, Prof. Lane Harrison suggested to look into IHME[1] website which contains the government policies of each state in U.S. and the enactment dates.

IHME focus on 6 policies which are

1. Mass gathering restrictions
2. Initial business closures
3. Education facilities closed
4. Non-essential services closed
5. Stay at home order
6. Travel severely limited



[1] IHME COVID-19 Projections

Data

From the IHME website, we extract 6 government policies and the enactment dates of each policy to the excel file.

Some of the policies in some states are not used so there are some NaN data. However, the “travel severely limited” policy is not being used in any states. Thus, we decided to not include it in our dataset.

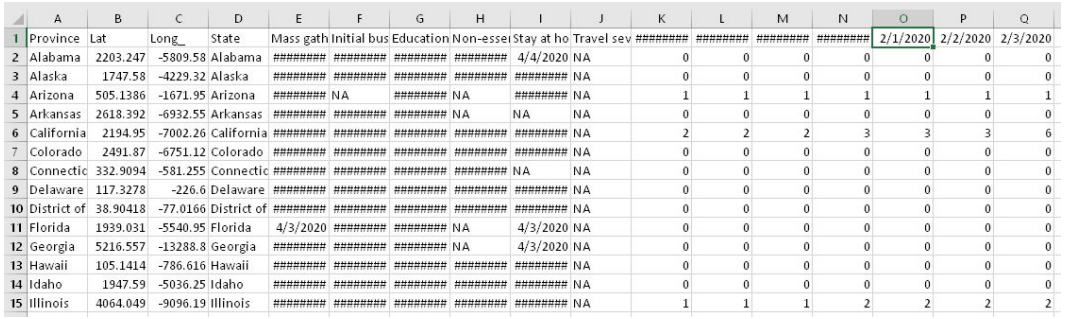
	A	B	C	D	E	F	G
1	State	Mass gathering restrictions	Initial business closure	Educational facilities closed	Non-essential services closed	Stay at home order	Travel severely limited
2	Alabama	3/19/2020	3/19/2020	3/19/2020	3/28/2020	4/4/2020	NA
3	Alaska	3/24/2020	3/17/2020	3/16/2020	3/28/2020	3/28/2020	NA
4	Arizona	3/30/2020	NA	3/16/2020	NA	3/30/2020	NA
5	Arkansas	3/27/2020	3/19/2020	3/17/2020	NA	NA	NA
6	California	3/11/2020	3/19/2020	3/19/2020	3/19/2020	3/19/2020	NA
7	Colorado	3/19/2020	3/17/2020	3/23/2020	3/26/2020	3/26/2020	NA
8	Connecticut	3/12/2020	3/16/2020	3/17/2020	3/23/2020	NA	NA
9	Delaware	3/16/2020	3/16/2020	3/16/2020	3/24/2020	3/24/2020	NA
10	District of Columbia	3/13/2020	3/16/2020	3/16/2020	3/25/2020	3/30/2020	NA
11	Florida	4/3/2020	3/17/2020	3/17/2020	NA	4/3/2020	NA
12	Georgia	3/24/2020	3/24/2020	3/18/2020	NA	4/3/2020	NA
13	Hawaii	3/17/2020	3/17/2020	3/19/2020	3/25/2020	3/25/2020	NA
14	Idaho	3/25/2020	3/25/2020	3/23/2020	3/25/2020	3/25/2020	NA
15	Illinois	3/13/2020	3/16/2020	3/17/2020	3/21/2020	3/21/2020	NA
16	Indiana	3/12/2020	3/16/2020	3/19/2020	3/24/2020	3/25/2020	NA
17	Iowa	3/17/2020	3/17/2020	4/4/2020	3/17/2020	NA	NA
18	Kansas	3/17/2020	NA	3/17/2020	NA	3/30/2020	NA
19	Kentucky	3/19/2020	3/16/2020	3/20/2020	3/26/2020	NA	NA
20	Louisiana	3/13/2020	3/17/2020	3/16/2020	3/22/2020	3/23/2020	NA
21	Maine	3/18/2020	3/18/2020	3/16/2020	3/25/2020	4/2/2020	NA
22	Maryland	3/16/2020	3/16/2020	3/16/2020	3/23/2020	3/30/2020	NA
23	Massachusetts	3/13/2020	3/17/2020	3/17/2020	3/24/2020	NA	NA
24	Michigan	3/13/2020	3/16/2020	3/16/2020	3/23/2020	3/24/2020	NA
25	Minnesota	3/28/2020	3/17/2020	3/18/2020	NA	3/28/2020	NA
26	Mississippi	3/24/2020	3/24/2020	3/19/2020	4/3/2020	4/3/2020	NA
27	Missouri	3/23/2020	3/23/2020	3/23/2020	NA	4/6/2020	NA

Data

For the JHU dataset, it is a county-level data. Thus, we filtered to get only U.S. states and combine with IHME policies dataset.

Finally, our final data has the state names, 5 main policies enactment time and number of confirmed cases in each day from Feb 1st to May 1st.

We then calculated our measure of variance for each policy in each state and add them into the final dataset.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	Province	Lat	Long_	State	Mass gath	Initial bus	Education	Non-esse	Stay at ho	Travel sev	########	########	########	########	2/1/2020	2/2/2020	2/3/2020
2	Alabama	2203.247	-5809.58	Alabama	########	########	########	########	4/4/2020	NA	0	0	0	0	0	0	0
3	Alaska	1747.58	-4229.32	Alaska	########	########	########	########	########	NA	0	0	0	0	0	0	0
4	Arizona	505.1386	-1671.95	Arizona	########	NA	########	NA	########	NA	1	1	1	1	1	1	1
5	Arkansas	2618.392	-6932.55	Arkansas	########	########	########	NA	NA	NA	0	0	0	0	0	0	0
6	California	2194.95	-7002.26	California	########	########	########	########	########	NA	2	2	2	3	3	3	6
7	Colorado	2491.87	-6751.12	Colorado	########	########	########	########	########	NA	0	0	0	0	0	0	0
8	Connectic	332.9094	-581.255	Connectic	########	########	########	########	NA	NA	0	0	0	0	0	0	0
9	Delaware	117.3278	-226.6	Delaware	########	########	########	########	########	NA	0	0	0	0	0	0	0
10	District of	38.90418	-77.0166	District of	########	########	########	########	########	NA	0	0	0	0	0	0	0
11	Florida	1939.031	-5540.95	Florida	4/3/2020	########	########	NA	4/3/2020	NA	0	0	0	0	0	0	0
12	Georgia	5216.557	-13288.8	Georgia	########	########	########	NA	4/3/2020	NA	0	0	0	0	0	0	0
13	Hawaii	105.1414	-786.616	Hawaii	########	########	########	########	########	NA	0	0	0	0	0	0	0
14	Idaho	1947.59	-5036.25	Idaho	########	########	########	########	########	NA	0	0	0	0	0	0	0
15	Illinois	4064.049	-9096.19	Illinois	########	########	########	########	########	NA	1	1	1	2	2	2	2

Exploration Data Analysis

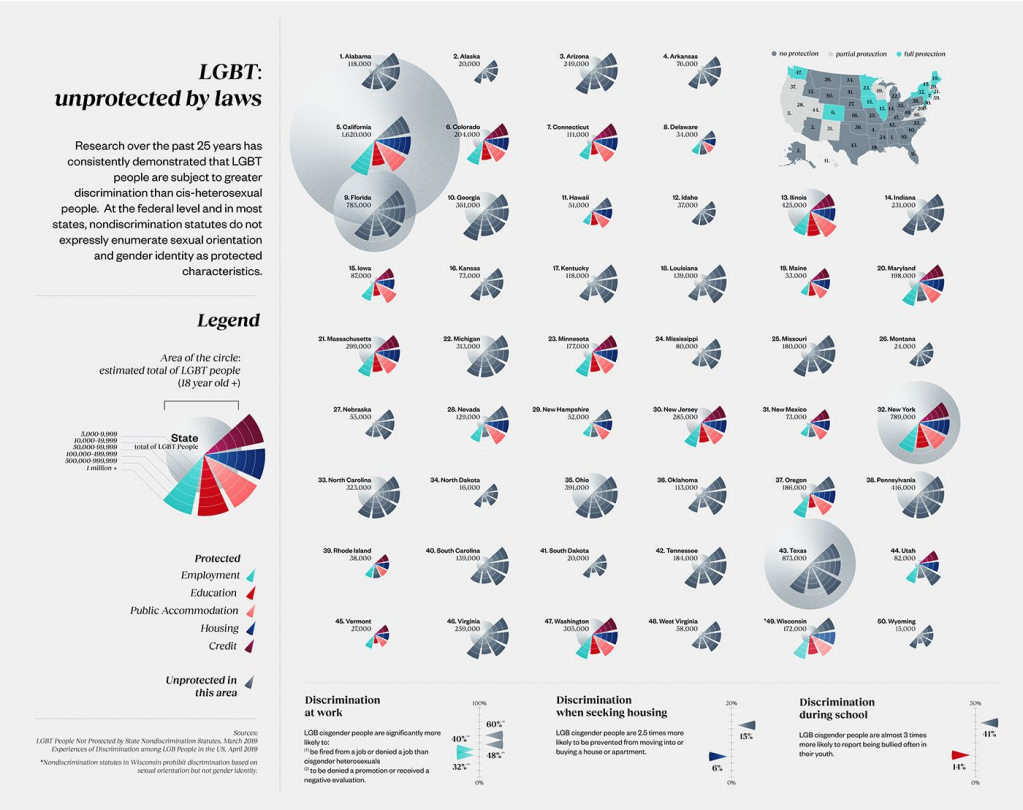
We explored JHU confirmed cases data using time trend visualization. As we expected the trend was ascending in all states over the time horizon. We realised just using time trend visualization would not be enough for us to achieve our goal and answer our question.

We decided to focus on two major indices in our data:

1. The date in which each policy was enacted in each state.
2. The value of measure of variance for each policy in each state, that is, the ratio of increase in number of confirmed cases, one week after enactment time of a policy in a state.

We used uniform and quantile distribution to understand the characteristic of measure of variance values for different policies and see how these values are distributed.

Design Evolution



Since we plan to work on each state in U.S., we look up to some infographics and found the LGBT: unprotected by laws visualization.

This visualization was the based-line for designing our project. We develop the same idea as this visualization by using the five policies as same as the five categories in the LGBT visualization.

We were aware that pie chart is not a good design as the angle is not high ranked magnitude chanel in terms of perception, however in this visualization the designer is not using the angel but the length of each slice to show the magnitude.

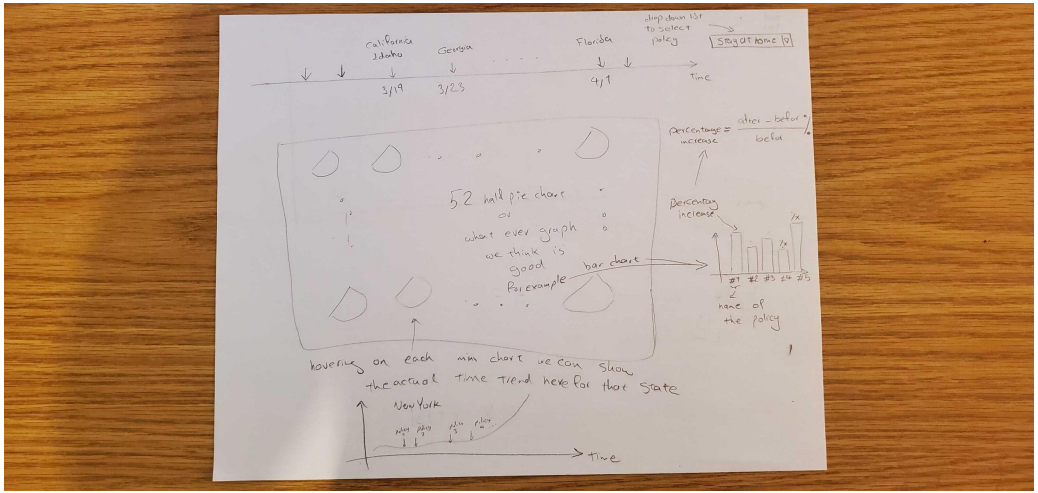
Design Evolution

Inspired by LGBT infographics, we draw our interface which contains half pie charts for each state.

At the top, we plan to make a timeline which shows the date of each policy.

At the bottom, we plan to make a total number of confirmed cases in time series graph of all states.

At the right, we put the equation about variance and some bar charts.

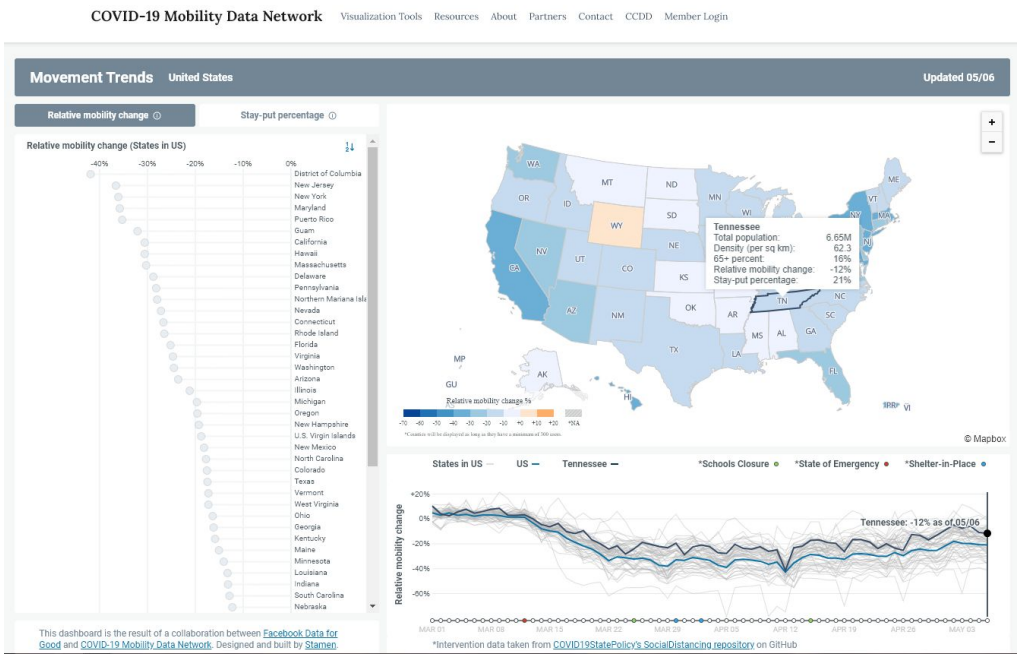


Design Evolution

Another visualization that inspired us was COVID-19 mobility data network[1]. The horizontal Lollipop chart at the left seems to be appropriate for us to sort out the states on one axis and show the enactment time of different policies in each state.

At the bottom of trend data located at right-bottom of canvas, the designer used connected 1-D grid of empty circles and lines to show the enactment time of policies in each state by filling the corresponding circle with a specific color.

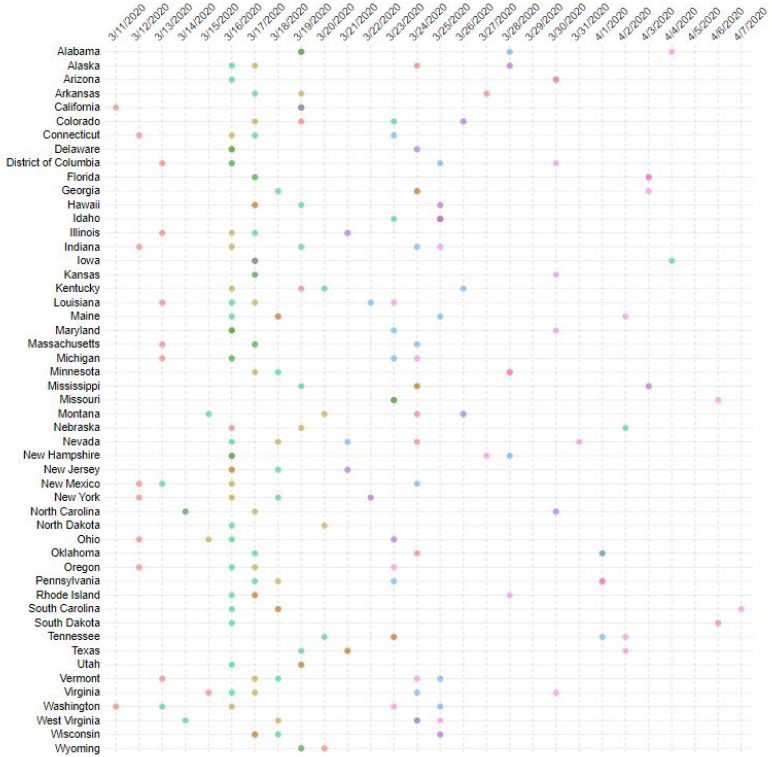
Combining these two, we based our design for visualizing the enactment times.



[1] COVID-19 Mobility Data Network

Implementation

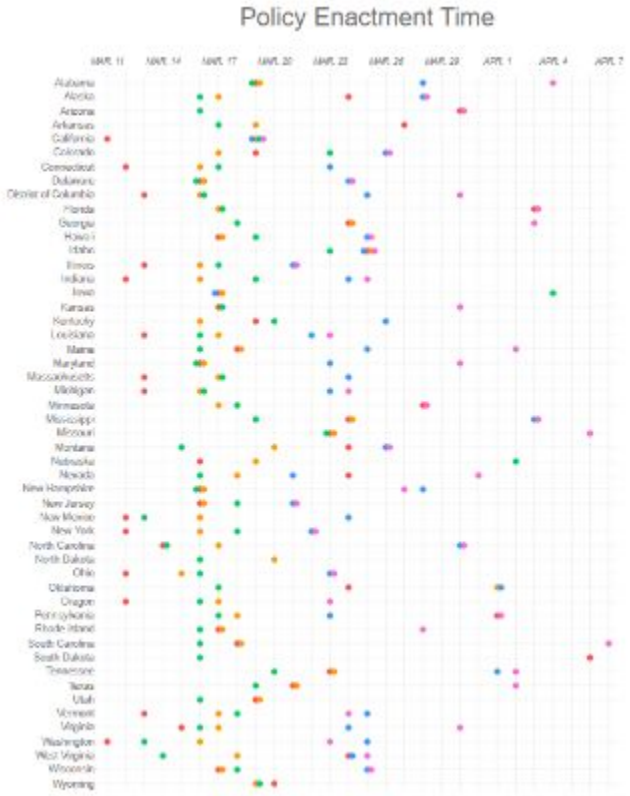
Policy Enactment Dates



We developed our first version of visualization as shown in the left picture. The horizontal axis is the Time and vertical axis is the name of states. Dots with different colors show the 5 policies and the positions of them in graph show the enactment time of each policy in each states.

Since some of the states do not use all the policies, there may be less than five dots for these states.

Implementation

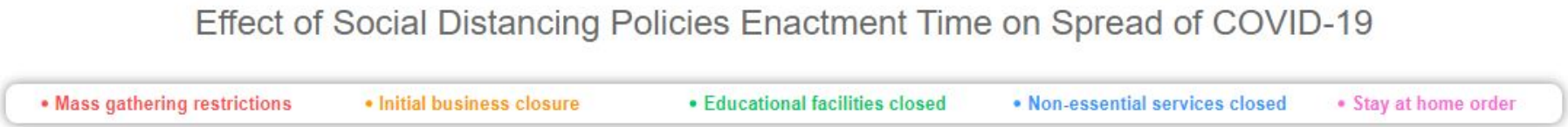


After creating the timeline visualization, we changed the tick date to interval of 3 days to get a clear date.

Then, we made a legend to indicate the policies as

- Mass gathering restrictions (red)
- Initial business closure (orange)
- Educational facilities closed (green)
- Non-essential services closed (blue)
- Stay at home order (purple)

Implementation



Before combining to the pie chart, we changed the format of legend. We also asked Prof. Harrison for an idea of legend to represent the color based on the context of each policy.

At the end, we thought our color choices were good enough to see the different between policies. Thus, we keep using the same colors with just a slight change.

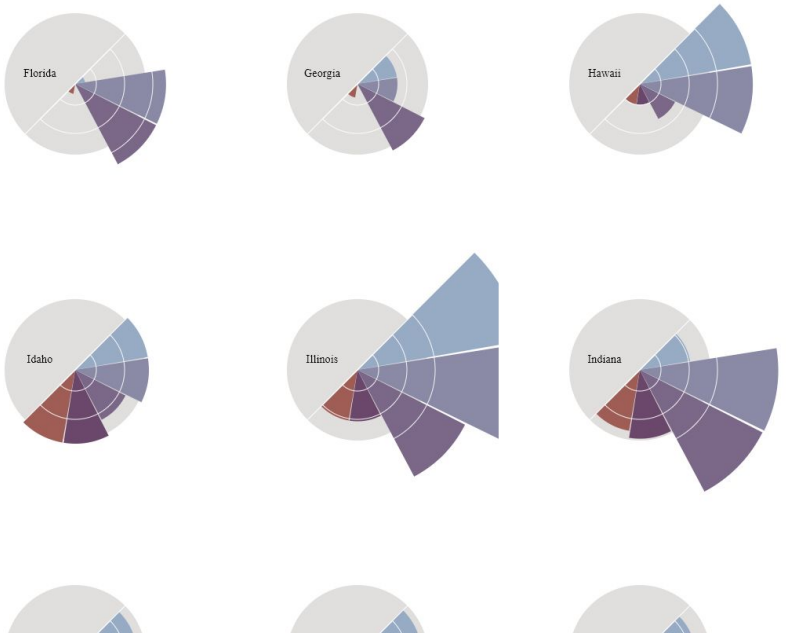
Implementation

Creating pie chart for each state using D3 function

- Arc : we use arc to create each policy by making the arc 36 angle with different position. We also use arc function to create the white indicate line because it is half of the circle.

- Circle : we use circle function to create the grey circle which indicate the total number of confirmed cases.

As we can see, we need to do the scaling.



Implementation

Since the variance of each policy is different, we have to create a function to assign a value to its interval.

For example, if the value is 12.34, it will be assigned to 3. The value 3 will be the radius of the arc.

We make six functions for five policies for arcs and the total cases for a circle.

Implementation

After we fixed the scaling problem, we change the color of the arcs to sync with the timeline visualization.

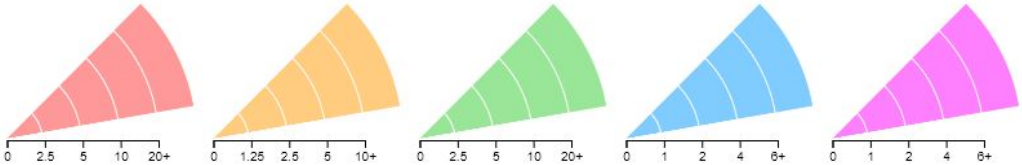
Then, we added mouseover and mouseout function to the arcs svg to create a text. When moving the cursor to the arc, it shows the policy number with the variance value.

If we move the cursor to the grey circle, it will show the total case number.

Implementation

Lastly for the pie chart, we made a legend to indicate the value of the arc. Since we match the color to the timeline legend, it is already indicated the policy. In here, we used the axis function to create the x-axis and set the number of tick to equal four.

We set the value in each policy differently because the variance is different.



Implementation

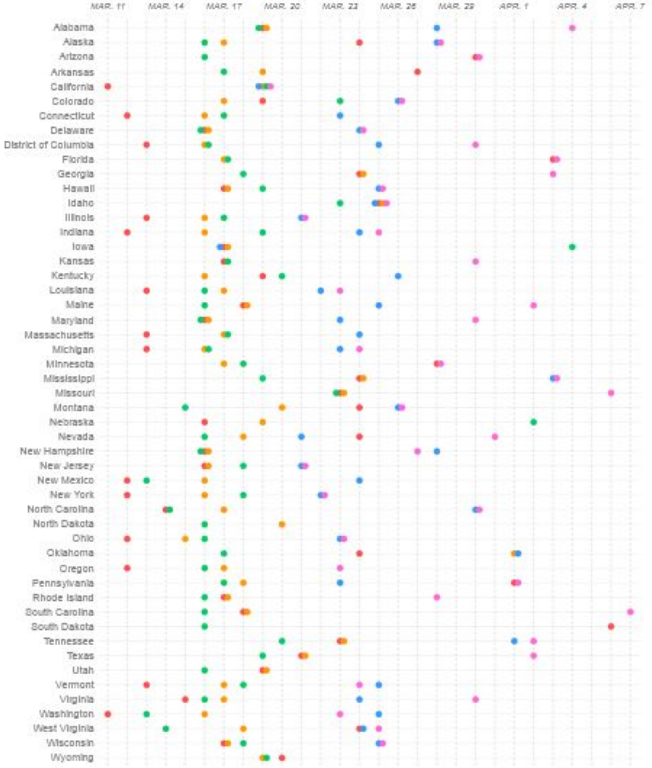
To finalize it, we combined both visualization together and made a header for our html. Then, we used the same idea in our multiple views assignment to map between two visualization.

Hovering on dots on the left panel, we change the opacity so that only the data associated with that policy pops up in the entire graph. In this way we can compare different states to see if enacting that particular policy earlier or later effected on the ratio of increase shown at right graph

Effect of Social Distancing Policies Enactment Time on Spread of COVID-19

• Mass gathering restrictions • Initial business closure • Educational facilities closed • Non-essential services closed • Stay at home order

Policy Enactment Time



Ratio of Increase, One Week After Enactment Time

Evaluation

By the Policy Enactment Time visualization, we can clearly see that most of the states used the mass gathering restrictions as the first policy to avoid the spreading. Later about a week, both initial business closure and education facilities closed policies were launched. And the last policy seemed to be stay at home order policy.

After exploring in the visualized data, we conclude that there is no significant effect, neither positive nor negative, between the enactment time of policies and the spread of the virus in different states.

This implies that more or less states took action at the right time, and other natural factors such as population, density or lifestyle in states, had the major effect on the spread of virus

For future work, we can change the after enactment period (e.g one month) to see if any significant correlation will be observed or not.