**The Impact of State Governor’s Political Party on the Reopening Policy and the Number of Covid-19 Cases in the US**

Nadia Fitriani

Yuwen Xiong

1. **Introduction/Executive Summary**

In this project, we aim to analyze the impact of reopening policy on the number of reported Covid-19 cases in the US by State during Mar 2020-Nov 2020. We want to see if State Governor’s party will have an impact on his/her decision of reopening policy and the number of new cases in the specific state.

The project contains two parts:

1. Web scraping text and downloading pre-existing datasets.
2. Do analysis and draw graphs to show results

Our datasets include:

1. Dataset for stay-at-home order expiration dates by state (web scraped 3 websites and tidied the data to ensure the accuracy)
2. Dataset for Governor’s Party by state (downloaded)
3. Dataset for Covid cases (downloaded and tidied)

Our final dataset used for follow-up analysis is “*data.csv*” in the folder, where we merged and tidied all three datasets above.

Our main finding is that…

The coding skills that we used are:

1. Web scraping (a lot & different structured websites)
2. Create informative and elaborate plots

The limitation of the project is…

1. **Results**
2. **Explanation of Working Process**
3. Web scraping text and downloading pre-existing datasets.

We scrapped text to get the date in which reopening policy was implemented. Here, we first need to define the reopening policy. Since each state has different policies/executive orders/phases, and [the definition of reopening](https://www.multistate.us/issues/covid-19-state-reopening-guide) varies a lot across people, it is hard to find a general rule that tells us the extent of reopen in one state. Thus, instead of looking at the reopening policy directly, we paid attention to the date when the stay-at-home order expires. The expiration of the stay-at-home order indicates that people are not strictly quarantined at home. Instead, they can go to public places and resume regular social activities while still remain 6-feet social distance. It indicates that the economy of the state is generally on the track of getting back to normal. Most states have stay-at-home order that expired from Apr to Jun. Thus, it is more convenient to do the comparison as most stay-at-home orders are similar to each other.

We checked three websites to get the accurate results for the expiration of the stay-at-home order: [HUSCHBLACKWELL](https://www.huschblackwell.com/state-by-state-covid-19-guidance), [NGA](https://www.nga.org/coronavirus-reopening-plans/) and [MULTISTATE](https://docs.google.com/document/u/1/d/e/2PACX-1vSXZCFCbIRiRDRC-SWyc36T0S0hjXxT9wZAGM4V01_xtbywLBEn0o_kgmfs0dMJ4VbpPh30j2ZFZ3TH/pub). We first tried Huschblackwell, as it has very detailed explanation by date about how the reopen progressed in each state. However, it is very complicated to get the accurate reopen date due to the diverse definition of reopen and the complicatedness of the text on this website. We tried to use NLP and keyword index, but found it is even hard to include all keywords that indicates reopen or stay-at-home order. Thus, we decided to jump to other sources. We find NGA but too many states’ dates are missing. Thus, we changed to Multistate. It is very clearly organized and relatively easy to scrape. It worked pretty well, though there are a few states that do not have data. We manually checked their official websites, news and information on Huschblackwell and input those dates.

In terms of data regarding Covid cases and Governor’s political party, there are pre-existing dataset on the internet. Thus, we downloaded and tidied the data. Then we merged the two datasets with the dates that we generated by web scraping.

1. Do analysis and draw graphs to show results
2. **Discussion**
3. **Conclusion**