COVID Death Rates and U.S. Politics

Introduction

In the United States, COVID-19 mitigation strategies have been implemented in a patchwork manner. States, counties, and towns across the U.S. used different variations of mask mandates, stay-at-home orders, and/or guidelines for private businesses on different timelines. Inconsistencies in policy details, compliance, and timelines make it difficult to directly compare the effectiveness of specific policies.

Broadly speaking, Republicans and Democrats have been divided in their response to COVID-19. Democratic officials tend to favor more aggressive mitigation strategies than their Republican counterparts. To investigate the outcomes of these differing approaches, this paper will compare regional election results with the corresponding per capita COVID-19 death rates. The margin of Republican or Democratic votes is assumed to be a proxy for local COVID-19 policy.

```
library(tidyverse)
library(ggplot2)
library(lubridate)
library(dataverse)
```

Datasets

This investigation requires county-level data on COVID-19 mortality rates and 2020 election results. COVID-19 mortality rates will be calculated with data provided by Johns Hopkins University. Precinct-level 2020 Presidential election data is provided by the MIT Election Data and Science Lab (hosted by Harvard Dataverse).

```
#COVID-19 data provided by the Johns Hopkins University Center for Systems Science and Engineering
url <- "https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/csse_covid_19_data/csse_covid_19_time_series/time_serie
data.us_deaths.raw <- read_csv(url, show_col_types = FALSE)
```

```
#2020 U.S. presidential election results
data.elections.raw <- get_dataframe_by_name(
    filename = "countypres_2000-2020.tab", dataset = "10.7910/DVN/VOQCHQ",original = FALSE,
    server = "dataverse.harvard.edu")
```

```
#function to estimate missing county vote totals (necessary to calculate vote margin)
fn.fill_na_totals <- function(candidatevotes, totalvotes){
  estimate = sum(candidatevotes)
  ifelse( is.na(totalvotes), estimate, totalvotes)
}

#aggregate county level votes
data.elections <- data.elections.raw %>%
  filter(year == 2020, candidate != "OTHER", candidate != "JO JORGENSEN") %>%
```

```
  group_by(state, county_name, candidate) %>%
  mutate(candidatevotes = sum(candidatevotes),
         county_fips = as.numeric(county_fips)) %>%
  select(!c(version, mode)) %>%
  unique.data.frame() %>%
  group_by(state, county_name) %>%
  mutate( totalvotes = sum(candidatevotes),
          pct_votes = round(candidatevotes/totalvotes *100, 2),
          vote_margin = 2*pct_votes - sum(pct_votes)) %>%
  ungroup()
```

Generally speaking, state governors hold the authority to enact policies to mitigate the spread of COVID-19. Therefore, we will also consider the impact of the state governor's political affiliation on COVID-19 death rates. Governor data comes from Civil Services – a nonpartisan, non-profit organization which aggregates data regarding elected officials in the United States.

```
#current state governors
url.gov <- "https://raw.githubusercontent.com/CivilServiceUSA/us-governors/a40d49614c18bfa8bedd19fc6687018fef42bc70/us-governors
data.gov<- read_csv(url.gov) %>%
  select(state_code, party) %>%
  mutate(party = as.factor(ifelse(party == "republican", "R", "D"))) %>%
  rename(state_po = state_code,
         gov_party = party)
```

We will use the presidential election voting margins as a proxy for a county's political leanings. The difference in vote share between the incumbent, President Trump, and his rival, President Biden, are calculated such that a positive value represents a margin in favor of Trump.

```
#code county votes by margin for/against the incumbent, President Trump.
data.joined <- data.elections %>%
  filter(party == "REPUBLICAN") %>%
  full_join(data.gov)

#filter data to exclude US territories
data.us_deaths <- data.us_deaths.raw %>%
  filter(FIPS <= 56045 & FIPS > 1000) %>%
  select(FIPS, Population, `9/18/21`) %>%
  rename(covid_deaths = `9/18/21`,
         county_fips= FIPS) %>%
  mutate(deaths_per_thou = covid_deaths/Population *1000)

#join data tables, excluding counties with missing election data or no COVID deaths reported
data.joined <- data.joined %>%
  full_join(data.us_deaths) %>%
  filter(deaths_per_thou > 0 &
        !is.na(candidatevotes))
```

U.S. Counties

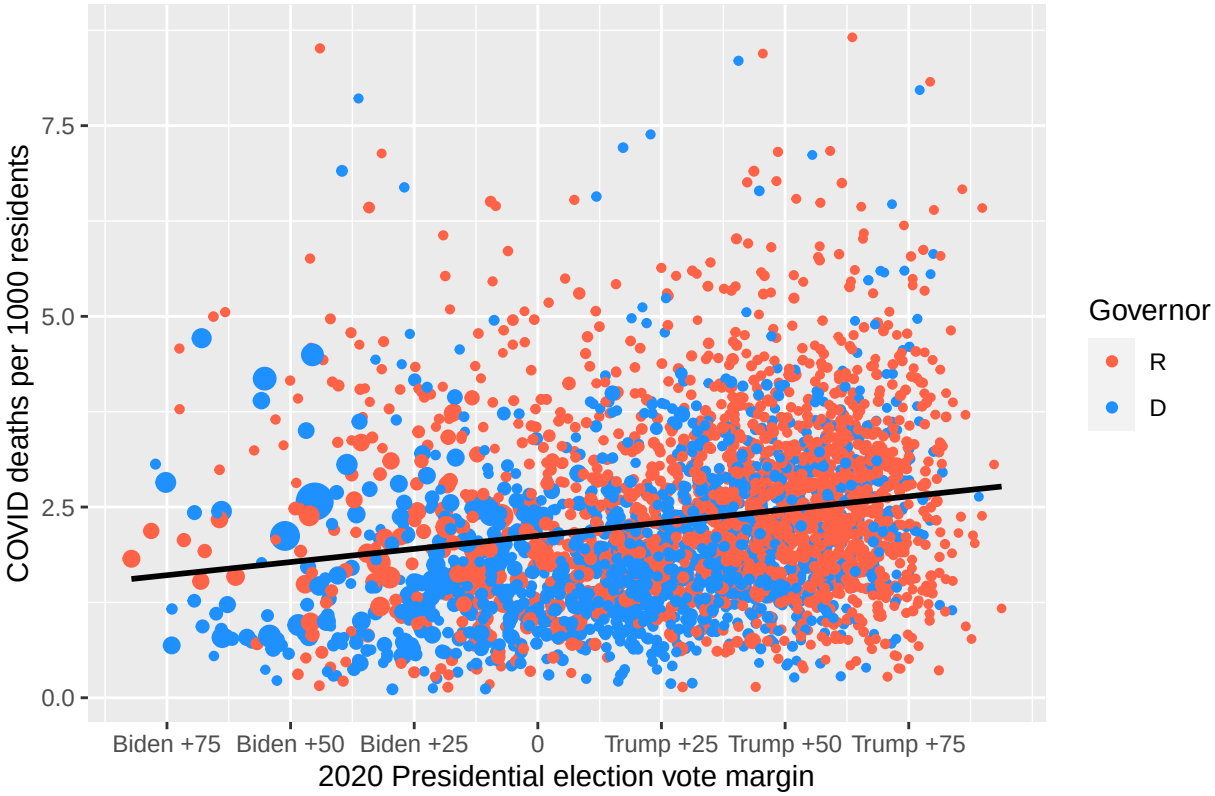
The following linear regression model and scatterplot indicate that there is a positive correlation between the degree to which a county supported President Trump and the county's per capita COVID-19 death rate. It is worth noting that the correlation is *very* weak (R-squared = .033)

```
model <- lm( data.joined$deaths_per_thou ~ data.joined$vote_margin )
summary(model)
```

```
##
## Call:
## lm(formula = data.joined$deaths_per_thou ~ data.joined$vote_margin)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.3350 -0.8403 -0.2127  0.5995  6.6939
##
## Coefficients:
##                           Estimate Std. Error t value Pr(>|t|)
## (Intercept)             2.1232161  0.0303569   69.94   <2e-16 ***
## data.joined$vote_margin 0.0068884  0.0006688   10.30   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.195 on 3063 degrees of freedom
## Multiple R-squared:  0.03348,    Adjusted R-squared:  0.03316
## F-statistic: 106.1 on 1 and 3063 DF,  p-value: < 2.2e-16
```

```
data.joined %>%
  mutate(pred = predict.lm(model, data.joined)) %>%
  ggplot(aes(x=vote_margin, y=deaths_per_thou)) +
  geom_point(aes(size=Population, color=gov_party)) +
  scale_color_manual(values= c("R" = "tomato", "D" = "dodgerblue")) +
  geom_line(aes(y=pred, size=2)) +
  labs(title="Figure 1: U.S. Counties",
       x = "2020 Presidential election vote margin", y = "COVID deaths per 1000 residents") +
  scale_x_continuous(breaks= seq(-75,75,25),
                     labels = c("Biden +75", "Biden +50", "Biden +25", "0", "Trump +25","Trump +50", "Trump +75")) +
  # theme(plot.caption = element_text(hjust=0)) +
  guides(size="none", color=guide_legend("Governor"))
```

Figure 1: U.S. Counties



U.S. Counties - Red states

When looking exclusively at the counties in "Red states" – states with a Republican governor – we see that there is no significant linear relationship (p-value = 0.11) between 2020 election margins and COVID-19 death rate. This may be evidence that COVID-19 policies enacted by Republican governors have a uniform impact throughout the state.

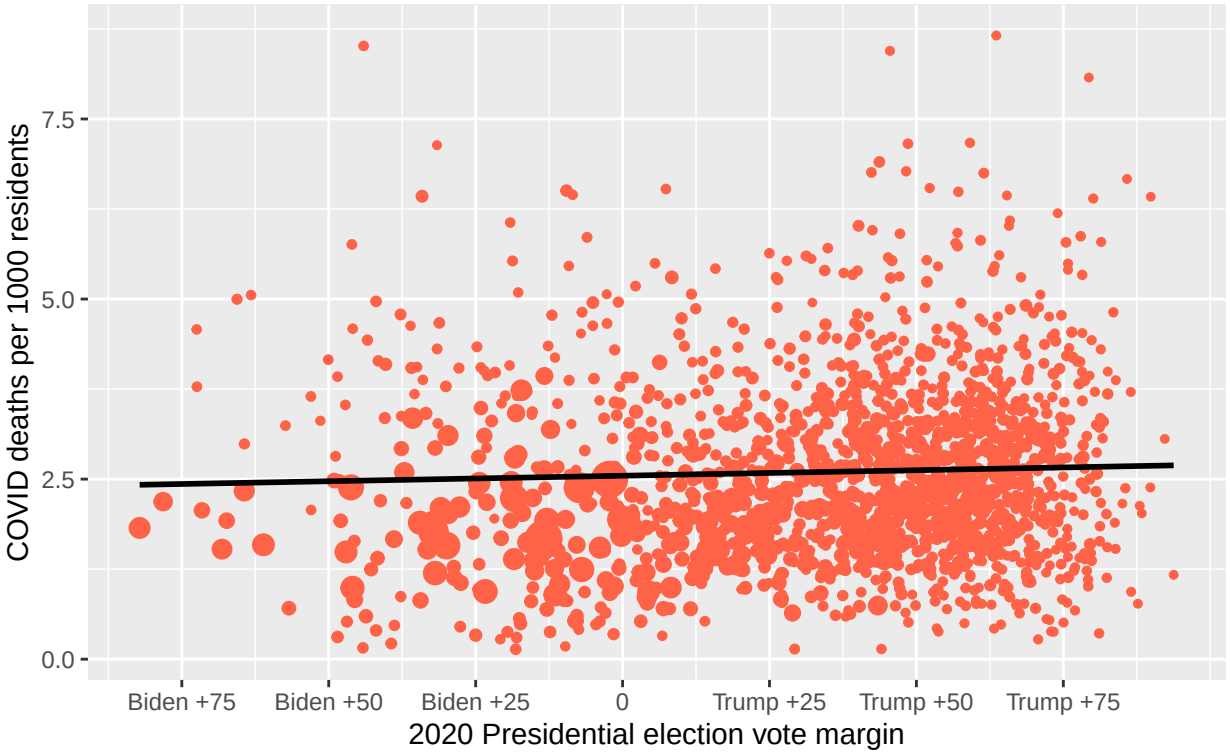
```
Rep.data <- data.joined %>%
  filter(gov_party == "R")

model.Rep <- lm( Rep.data$deaths_per_thou ~ Rep.data$vote_margin )
summary(model.Rep)
```

```
##
## Call:
## lm(formula = Rep.data$deaths_per_thou ~ Rep.data$vote_margin)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.4726 -0.8659 -0.1584  0.6758  6.0341
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)           2.547063   0.047504  53.618   <2e-16 ***
## Rep.data$vote_margin 0.001532   0.000960   1.595    0.111
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.239 on 1717 degrees of freedom
## Multiple R-squared:  0.00148,    Adjusted R-squared:  0.0008988
## F-statistic: 2.545 on 1 and 1717 DF,  p-value: 0.1108
```

```
Rep.data %>%
  mutate(pred.Rep = predict.lm(model.Rep, Rep.data)) %>%
  ggplot(aes(x=vote_margin, y=deaths_per_thou)) +
  geom_point(aes(size=Population, color=gov_party)) +
  labs(title="Figure 2: Red State counties",
       subtitle = "Counties under Republican governor from 2020-2021",
       x = "2020 Presidential election vote margin", y = "COVID deaths per 1000 residents") +
  scale_color_manual(values= c("R" = "tomato")) +
  geom_line(aes(y=pred.Rep, size=2)) +
  scale_x_continuous(breaks= seq(-75,75,25),
                     labels = c("Biden +75", "Biden +50", "Biden +25", "0", "Trump +25","Trump +50", "Trump +75")) +
  theme(legend.position="none", plot.caption = element_text(hjust=0))
```

Figure 2: Red State counties

Counties under Republican governor from 2020−2021



U.S. Counties - Blue states

When looking exclusively at the counties in "Blue states" – states with a Democratic governor – We see that there is a weak linear correlation (R-squared = .073) between election margins and COVID-19 mortality rates. This may be evidence that the efficacy of COVID-19 policies enacted by Democratic governors is contingent upon local factors.

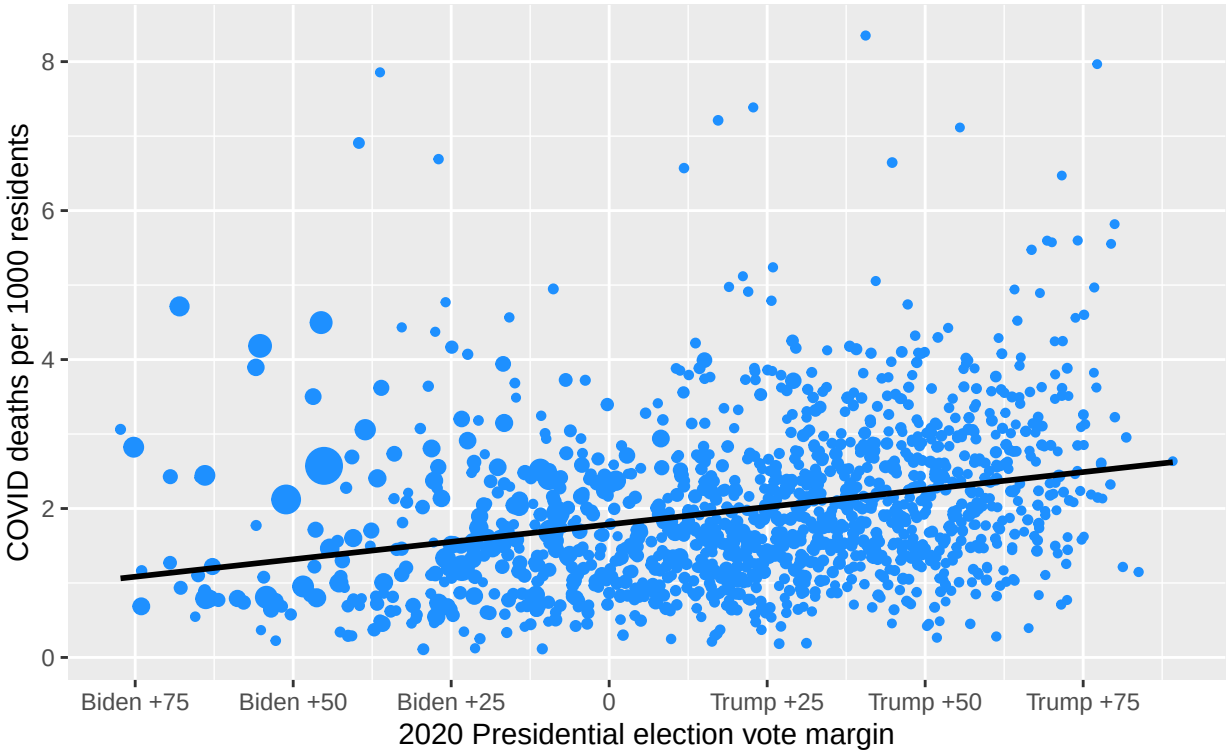
```
Dem.data <- data.joined %>%
  filter(gov_party == "D")

model.Dem <- lm( Dem.data$deaths_per_thou ~ Dem.data$vote_margin )
summary(model.Dem)
```

```
##
## Call:
## lm(formula = Dem.data$deaths_per_thou ~ Dem.data$vote_margin)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.0790 -0.7139 -0.1894  0.5157  6.4095
##
## Coefficients:
##                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.7865244  0.0357709   49.94   <2e-16 ***
## Dem.data$vote_margin 0.0093668  0.0009042   10.36   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.053 on 1344 degrees of freedom
## Multiple R-squared:  0.07394,    Adjusted R-squared:  0.07325
## F-statistic: 107.3 on 1 and 1344 DF,  p-value: < 2.2e-16
```

```
Dem.data %>%
  mutate(pred.Dem = predict.lm(model.Dem, Dem.data)) %>%
  ggplot(aes(x=vote_margin, y=deaths_per_thou)) +
  geom_point(aes(size=Population, color=gov_party)) +
  labs(title="Figure 3: Blue State counties",
       subtitle = "Counties under Democratic governor from 2020-2021",
       x = "2020 Presidential election vote margin", y = "COVID deaths per 1000 residents") +
  scale_color_manual(values= c("D" = "dodgerblue")) +
  geom_line(aes(y=pred.Dem, size=2)) +
  scale_x_continuous(breaks= seq(-75,75,25),
                     labels = c("Biden +75", "Biden +50", "Biden +25", "0", "Trump +25","Trump +50", "Trump +75")) +
  theme(legend.position="none", plot.caption = element_text(hjust=0))
```

Figure 3: Blue State counties

Counties under Democratic governor from 2020−2021

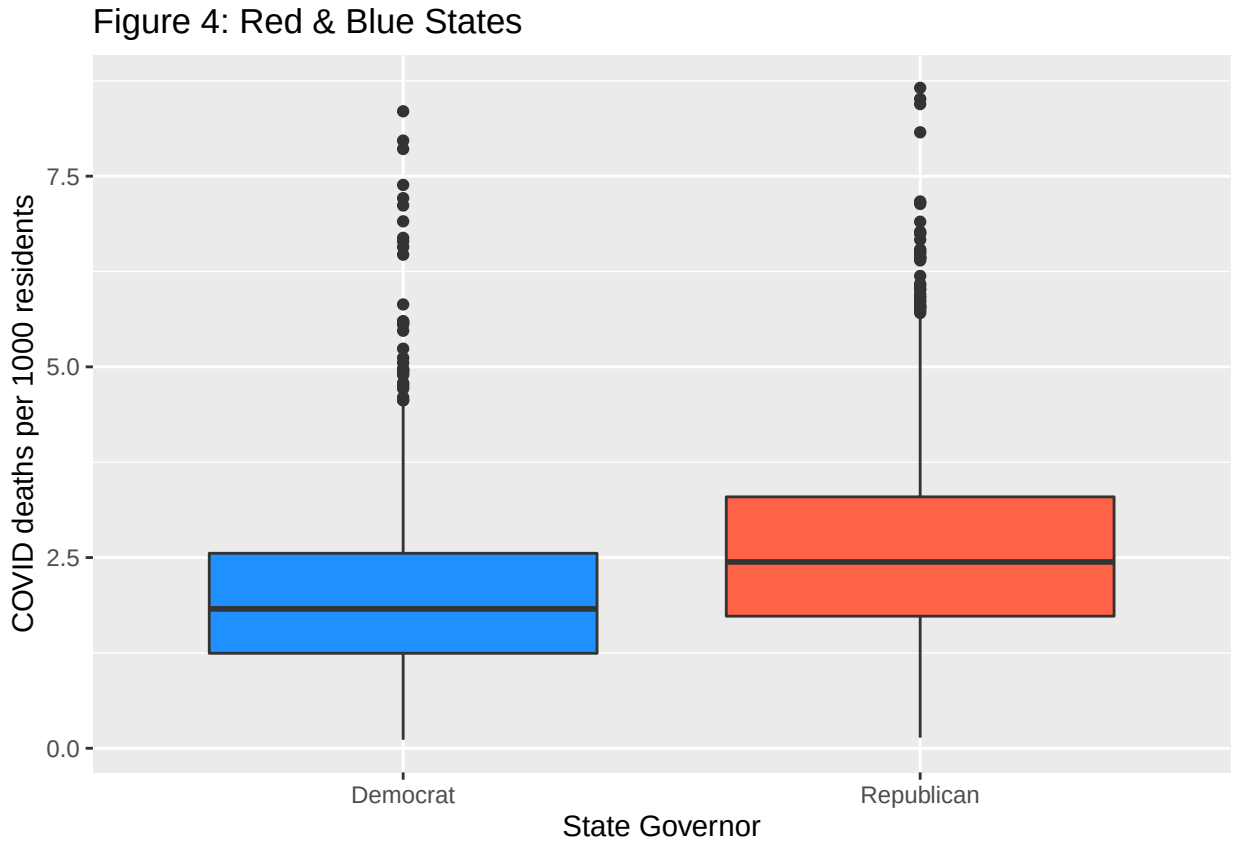


Comparing Red and Blue States

Interestingly, there appears to be a noteworthy difference in the correlation between vote margins and mortality rates in our "Red" and "Blue" states. The "Red" states (Republican governor) do not exhibit any linear correlation (p-value = 0.11) between vote margin and COVID-19 death rates. "Blue" states (Democratic governor) see a positive correlation between vote margin and COVID-19 death rates (R-squared = .073). In a Blue state, the model predicts that an extremely partisan Republican-leaning county would have double the mortality rate of an extremely partisan Democratic-leaning county.

To get a fuller picture of the relationship, we'll compare the means of Red and Blue state COVID-19 mortality rates. As demonstrated in the following t-test, the mean mortality rate is significantly different.

```
#gov_party boxplot
data.joined %>%
  ggplot(aes(x=gov_party, y=deaths_per_thou, fill = gov_party)) +
  geom_boxplot() +
  labs(title = "Figure 4: Red & Blue States",
       x = "State Governor",
       y = "COVID deaths per 1000 residents") +
  scale_x_discrete(labels=c("R" = "Republican", "D" = "Democrat")) +
  scale_fill_manual(values= c("R" = "tomato", "D" = "dodgerblue")) +
  theme(legend.position="none")
```




```
t.test(data.joined[data.joined$gov_party == "D",]$deaths_per_thou,
       data.joined[data.joined$gov_party == "R",]$deaths_per_thou)
```

```
##
##  Welch Two Sample t-test
##
## data:  data.joined[data.joined$gov_party == "D", ]$deaths_per_thou and data.joined[data.joined$gov_party == "R", ]$deaths_per
## t = -14.168, df = 3019.6, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.6812746 -0.5156318
## sample estimates:
## mean of x mean of y
##  2.007509  2.605963
```

Conclusion

This investigation concludes that the political landscape of the United States has a significant impact on regional COVID-19 mortality rates. Blue states see a weak but significant correlation between county-level political preference and mortality rate. Red states' mean county-level mortality rate is estimated to be 30% greater than in Blue states. Taken at face value, these results imply that the COVID-19 pandemic has been less deadly in areas governed by Democratic politicians. However, this investigation has numerous limitations to keep in mind.

Limitations:
In this investigation, we did not control for differences in the size, population density, demographics, or geography of counties. Therefore, it is possible that these factors underlie the relationship between regional political preference and COVID-19 death rates. Additionally, this investigation did not directly assess COVID-19 policies and did not enumerate the political composition of local officials such as mayors, state legislatures, or city council members who hold the power to implement and enforce local ordinances.

Biases: At this point in the COVID-19 pandemic, everyone has developed strong feelings about what mitigative policies should or should not be implemented. Although I have tried to keep any personal biases from influencing this investigation, it is pertinent to disclose that I believe vaccinations and mask mandates are effective tools for lessening the spread of infectious respiratory diseases.