

Language used by US Congress in Emails About COVID-19

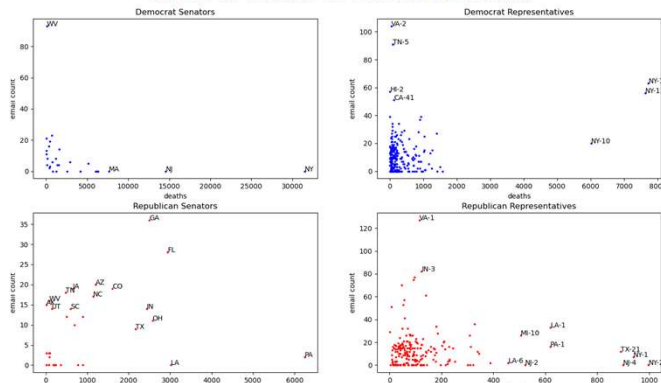
Kirsten Meidlinger and Lindsey Cormack, Ph.D



Pinnacle Scholars Summer
Research Program 2020

Lawmakers in congress are tasked with making decisions about the best ways to proceed during the global COVID-19 pandemic. The way congresspeople talk about COVID-19 is very important in this decision-making process. The language used when discussing the pandemic influences how decisions are made. We analyze congressional e-newsletters and whether those e-newsletters depended on each congressperson's local COVID-19 situation.

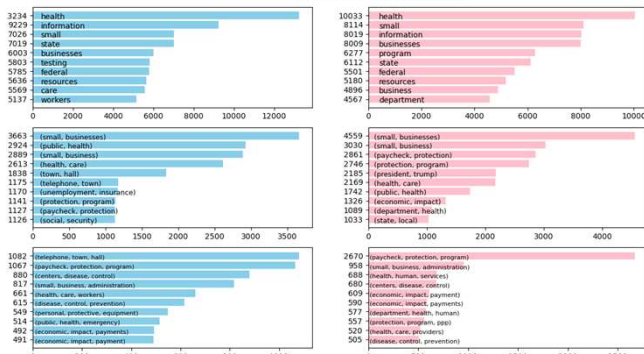
COVID-19 Deaths vs. E-newsletters sent



COVID-19 by District

I created a database of COVID-19 cases/deaths by congressional district. We had only seen COVID-19 data by state or by county, and congressional districts are sometimes completely separate from county lines. I used proportions of county populations and COVID-19 case/death data by county to create a database of estimated COVID-19 cases and deaths. We needed this data to see whether or not a relationship existed between amount of e-newsletters sent and local COVID-19 cases/deaths.

Most Common Phrases



Results

Despite the fact that states' reactions to the COVID-19 pandemic have been judged separately based on party, most federal legislators have in fact been spreading similar information and making similar arguments. There are not many stark differences between the language used by republican and democratic congresspeople in COVID-19 related e-newsletters.

Furthermore, COVID-19 case/death rates are not indicative of how many e-newsletters are sent out. While some states with large death counts have sent a lot of e-newsletters, other states with similar rates of death due to COVID-19 have not sent very many e-newsletters about COVID-19.

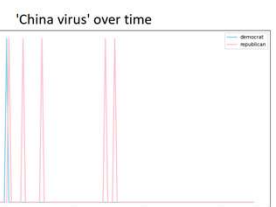
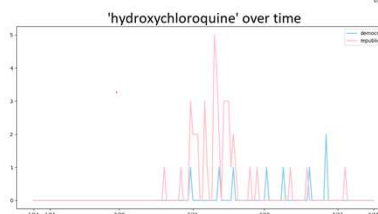
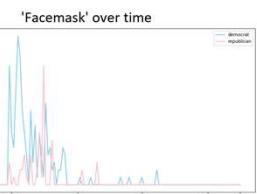
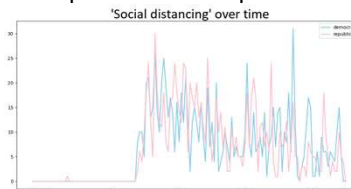
Hypotheses

- Legislators who see greater cases and deaths in their own districts send more COVID-19 e-newsletters.
- Republicans will refer to COVID-19 as 'China virus'/'Wuhan virus' more often than democrats.
- Republicans will talk about vaccines and hydroxychloroquine more frequently than democrats.
- Democrats will start talking about masks and other CDC safety guidelines sooner than republicans.

Language Analysis

For each email, I tracked the frequencies of words used, as well as two-word and three-word phrases (bigrams and trigrams) in each email. This way, I could have data for the most popular words and phrases used by democrats and republicans. To investigate our hypotheses, I built two email language analysis tools using Natural Language Toolkit (NLTK) in Python.

- A tool that shows how often a word or phrase is used by democrats and republicans in their emails over time.
- A tool that finds the most used words and phrases by party and one that displays usage by party trendlines of a specific word or phrase.



* Each mention of 'China virus' by republicans was plainly using the term to describe COVID-19, all but one mention by democrats was explaining the racist nature of the term