

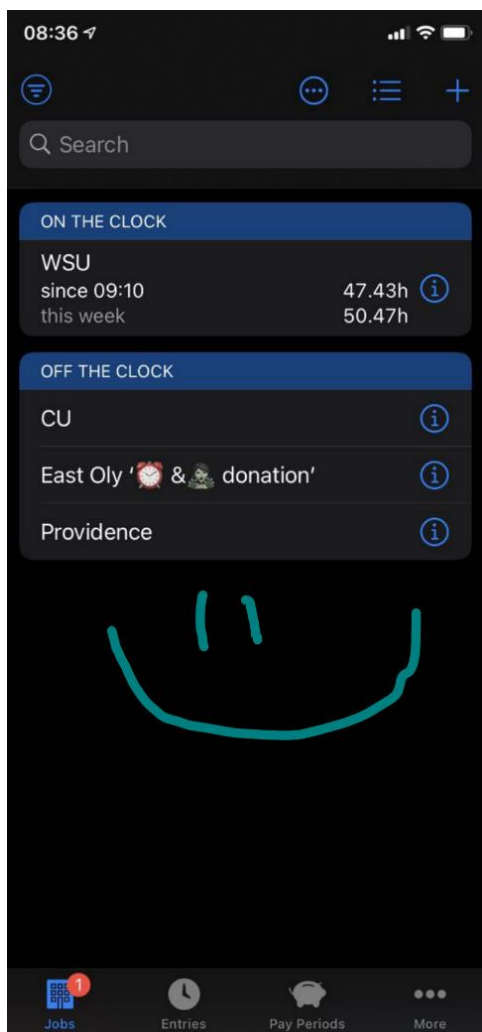
Well, due to unforeseen circumstances, I will not be able to submit my jupyter notebook today before 10am. I managed to break anaconda completely, and I cannot find anything that I thought was saved on my computer, but I am going to guess that when it deleted, it took all the files that it had in the contents with it. So, yay me?

However, I just uploaded 194 different files that were all the data that I had collected, scrubbed, updated, and had ready for this analysis. So, I guess I will just talk about what I did and why. So I apologize that my pretty choropleth is not going to be seen. However, I may just make it again later as it was fun, and looked awesome.

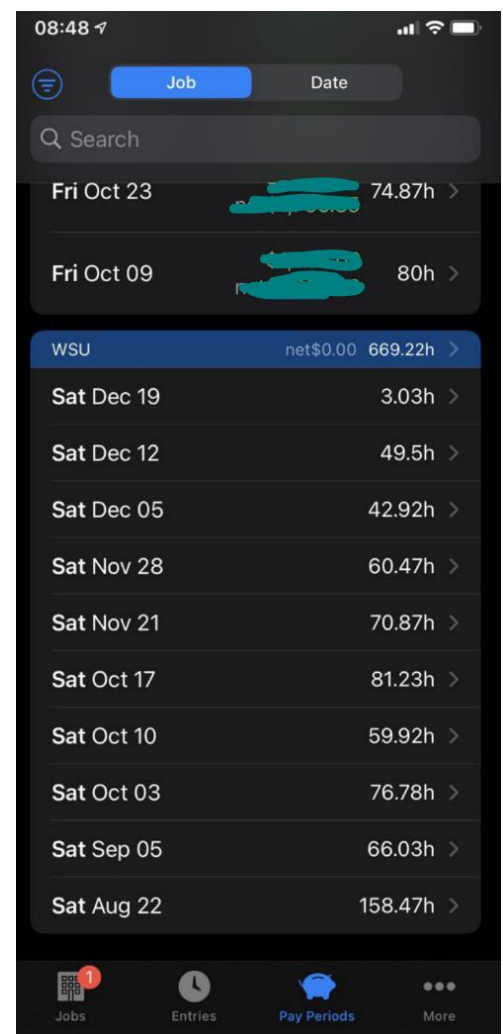
```
--no-color          Suppress colored output
--no-python-version-warning  Silence deprecation warnings for upcoming unsupported Pythons.

(base) C:\Users\sasha>conda install -c conda-forge geopandas
Collecting package metadata (current repodata.json): done
Solving environment: failed with initial frozen solve. Retrying with flexible solve.
Solving environment: failed with repodata from current repodata.json, will retry with next repodata source.
Collecting package metadata (repodata.json): done
Solving environment: failed with initial frozen solve. Retrying with flexible solve.
Solving environment: /
Found conflicts! Looking for incompatible packages.
This can take several minutes. Press CTRL-C to abort.
Examining json5: 9%
Examining zope.interface: 18%
Examining sphinxcontrib-devhelp: 26%
Examining gevent: 28%
Examining nltk: 28%
Examining conda-verify: 40%
Examining gmpy2: 41%
Examining pyflakes: 45%
Examining traitlets: 58%
Examining conda-package-handling: 59%
Examining sympy: 62%
Examining boto: 79%
Examining bottleneck: 92%
Examining jupyter-client: 93%
Examining jedi: 94%
Examining watchdog: 95%
Examining yapf: 96%
Examining conflict for prompt_toolkit pytest intervaltree threadpoolctl win_inet_pton pathlib2 networkx pywinpty sphinx/
Examining conflict for prompt_toolkit pytest intervaltree tbb libarchive win_inet_pton astroid pathlib2 networkx pywinp-
Examining conflict for pytest win_inet_pton pathlib2 networkx pywinpty xlwings pillow bokeh wincertstore python-libarch-
Examining conflict for blas m2w64-gcc-libc-core scipy anaconda m2w64-libwinpthread-git m2w64-gcc-libgfortran m2w64-gcc-
Examining conflict for pillow anaconda-navigator anaconda matplotlib-base bokeh libtiff imageio scikit-image: 13%
Examining conflict for blosc pillow zstd anaconda geopandas libtiff python-libarchive-c pytables libarchive conda-pac-
```

BRUH it.




I also know that it probably looks like I waited until the last minute to get this project in, however, that is not actually the case. It ended up being harder than I had expected to find some of the data for my analysis. So below are 2 screen shots. One of the total time I have spent working on this class alone. And the other is how long I have been working on this class in the last 2 days. I will give you a hint. I literally have been sitting at my desk for pretty much 2 days straight, getting up to go pee and once to shower really quick. So, with, over 650 hours of time dedicated to just this class alone.... Yeah, not really last minute.



HOW TO LIE WITH STATISTICS

Darrell Huff
Illustrated by Irving Geis



**Over Half a Million Copies Sold—
An Honest-to-Goodness Bestseller**

Covid cases are on the rise. They are still on the rise, and there are so many fallacies that spread around like wildfire in this world of misinformation which makes it really hard to know what is what. With that known, the covid cases that are spiking have so many variables to it, that it is never going to be as simple as "republican or democrat"

[illegible]

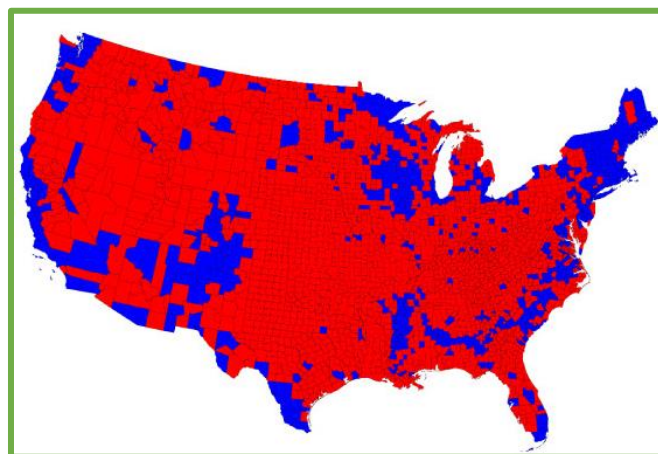
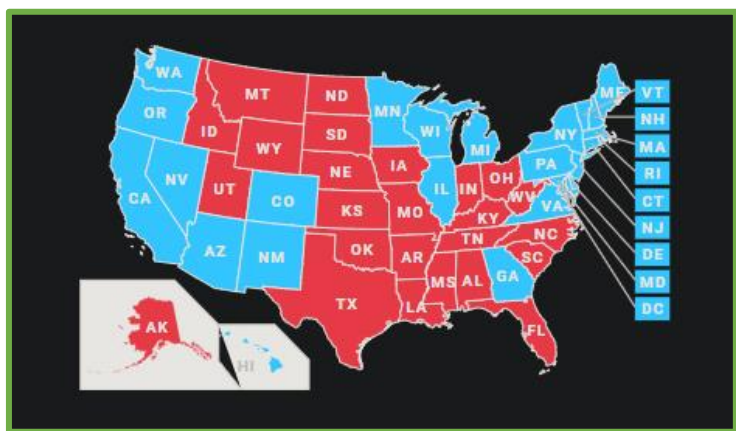
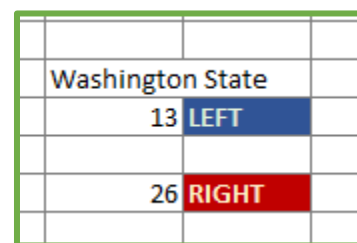
their article and premise that they had written about, I know that trying to look at it that way is not now an honest look at the data and the information, just as the way that their analysis was not honest or accurate.

As the people in one area do not live in a vacuum, we need to look at the data just as that. Each county, and each area has a mixture of both Left- and Right-wing individuals. Which mean that in each area, we would have a majority populace, and a minority. And everyone of those people are going to be somewhere on that political spectrum. (and this is where the choropleth that I had ready would have been amazing to see) but instead, I will just use the values that I had in my excel file:

987	Washington	Spokane	53063	20388	300	148576	135765	293294	50.66%	46.29%	96.95%	RIGHT	1764.0	296.4	522798
988	Washington	Stevens	53065	853	10	19808	7839	28283	70.04%	27.72%	97.75%	RIGHT	2478.0	18.5	45723
989	Washington	Thurston	53067	3776	50	65277	96608	166965	39.10%	57.86%	96.96%	LEFT	722.0	402.4	290536
990	Washington	Wahkiakum	53069	43	2	1741	1165	2972	58.58%	39.20%	97.78%	RIGHT	264.0	17.0	4488
991	Washington	Walla Walla	53071	2633	28	16400	13690	31055	52.81%	44.08%	96.89%	RIGHT	1270.0	47.8	60760
992	Washington	Whatcom	53073	2812	61	50489	83660	137660	36.68%	60.77%	97.45%	LEFT	2107.0	108.8	229247
993	Washington	Whitman	53075	2511	25	9067	11184	20964	43.25%	53.35%	96.60%	LEFT	2159.0	23.2	50104
994	Washington	Yakima	53077	16852	309	50555	43179	95713	52.82%	45.11%	97.93%	RIGHT	4296.0	58.4	250873
995	West Virginia	Barbour	54001	543	5	5116	1457	6677	76.62%	21.82%	98.44%	RIGHT	341.0	48.2	16441
996	West Virginia	Berkeley	54003	4224	36	33279	17186	51535	64.58%	33.35%	97.92%	RIGHT	321.0	371.2	119171

I live in Olympia. We are Thurston county. WSU is in Whitman county. Both of our counties are a majority of people that one could assume that they lean left in the political world. However, here in Thurston, of the 166,965 voters, there were still 65,277 people in this area that still vote right-wing. Lumping us all together as one blue state as was done in the previous article. And Washington, just like the rest of the US has a majority of right-wing counties in comparison to the left-wing counties.

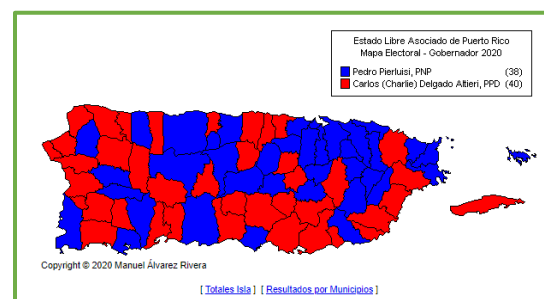
To be exact, in county majority rule, they have the left-wing voters beat 2:1. And despite this map being the one that a majority of people look at when they consider American politics and voting:



(minus the blue Georgia- because I don't think anyone really saw that one coming)

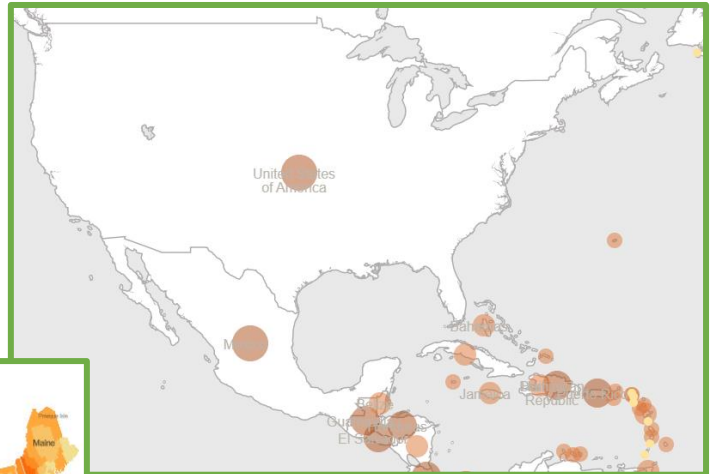
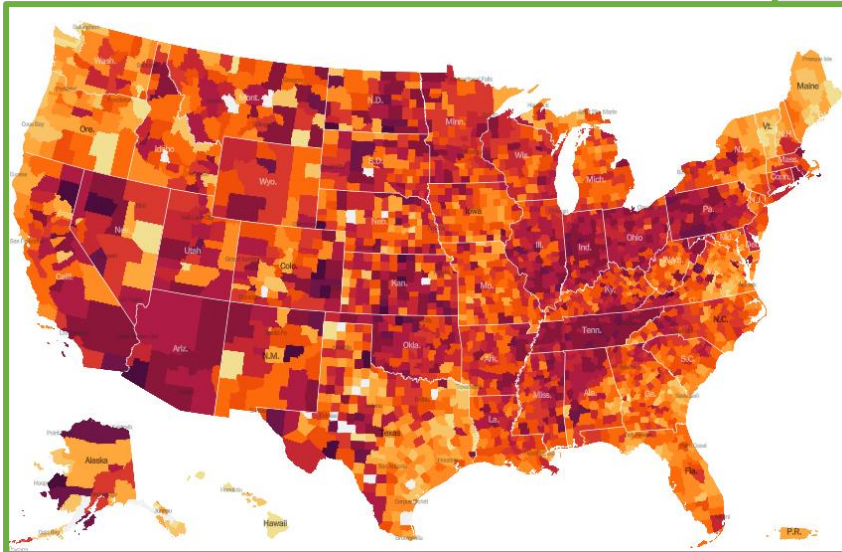
But the map on the right is a more accurate depiction of where the voters lie, and the striking color contrast that looks nothing like the map on the other side. (the one I made looked prettier, and had Puerto Rico, and the Virgin Islands as well) but for now, I guess this will do.

So since this virus spreads at an alarmingly high rate, what other reason could there be for the hot spots that are all across the US? What other argument could have been used for that article, rather than bad skewed political data? Because as much as I would love to be able to point a finger at someone, there is just not sufficient data for the claim of it being solely political.



Well, the map that is shown to us by the WHO is not the most helpful. However, it does show quite a few cases down around the Virgin Islands. But we will come back to that one a moment.

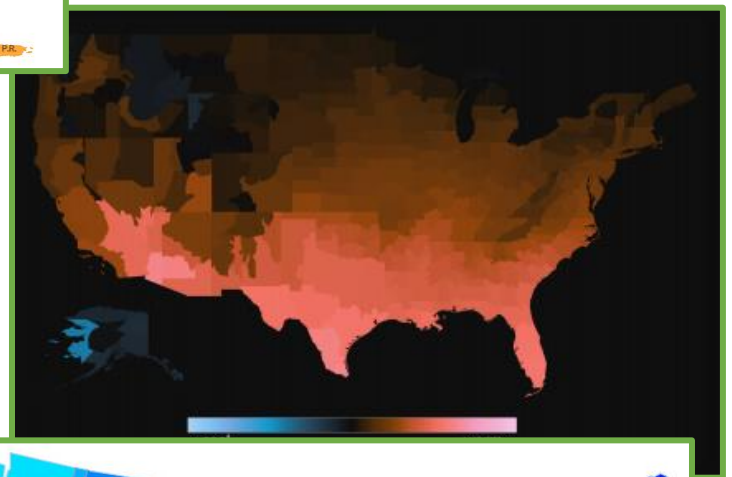
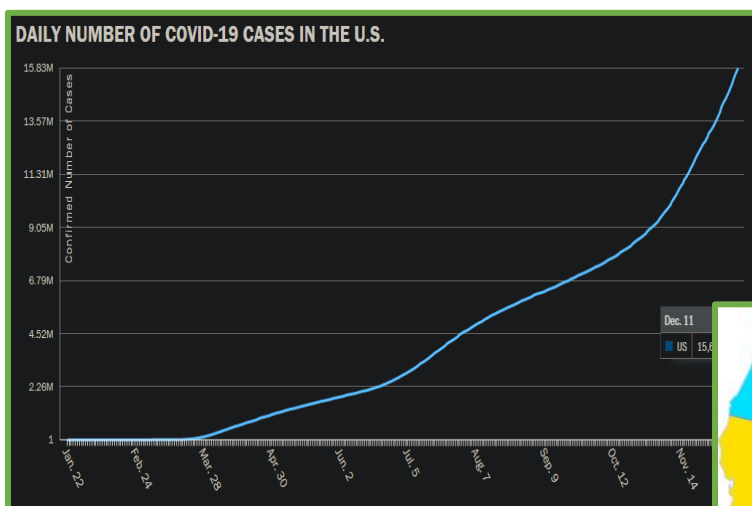
When I was looking up all the information for all the counties, I ran across the ce.naco.org website. And if I take the few seconds to turn off my Dark Reader, we get a map of covid cases that looks like this:



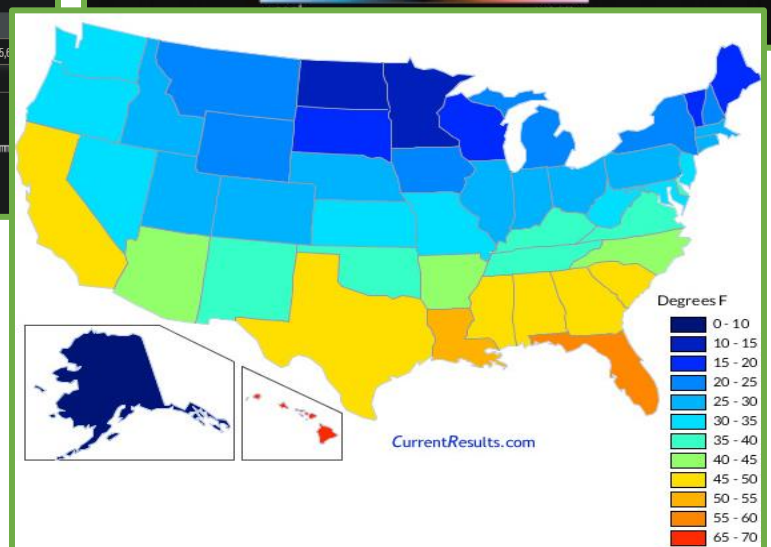
We will also get back this one in a second as well. I just wanted to point out first that the secondary conclusion by the article that we read came to the idea that there was a strong correlation with weather and the increase in cases, as was shown by many of their graphs:

here is the case rise, set for the same

day when I refreshed the data for my project:



And the average temperatures for the US. Although, yes there are more cases in the warmer states, but that does not account for the areas that are not warm, and are also having spikes in the cases, and the deaths in the united states.

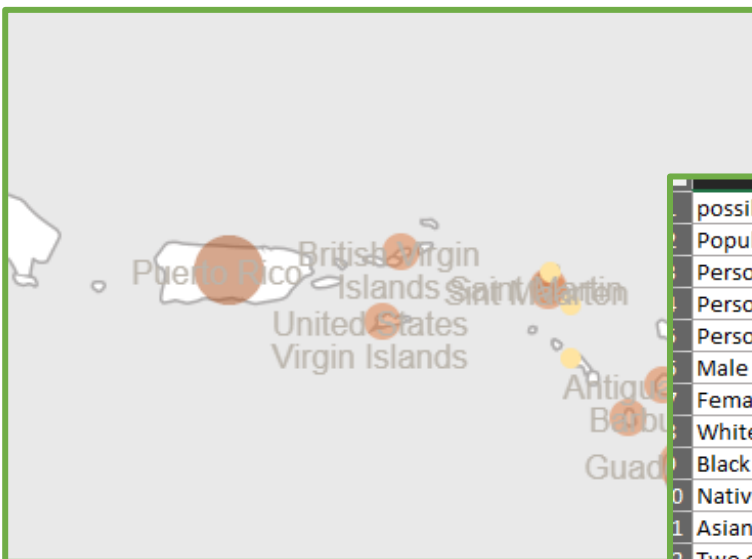


Another option that we could also look at is, besides a majority of those states having a high right-wing voter majority, is something else all those states have in common? California, Florida, North Dakota, Colorado, New York, Tennessee... pretty much everywhere, minus Hawaii who is trying to stay out of this mess. I had pulled the data for a few things that may also be a contributing factor to see if there was any correlation. These were also on the notebook. = [however, I do have the excels at least:

Most of the information that was on here did not really have any kind of strong correlation with the current infection rates. One area did have a higher correlation than the others did.

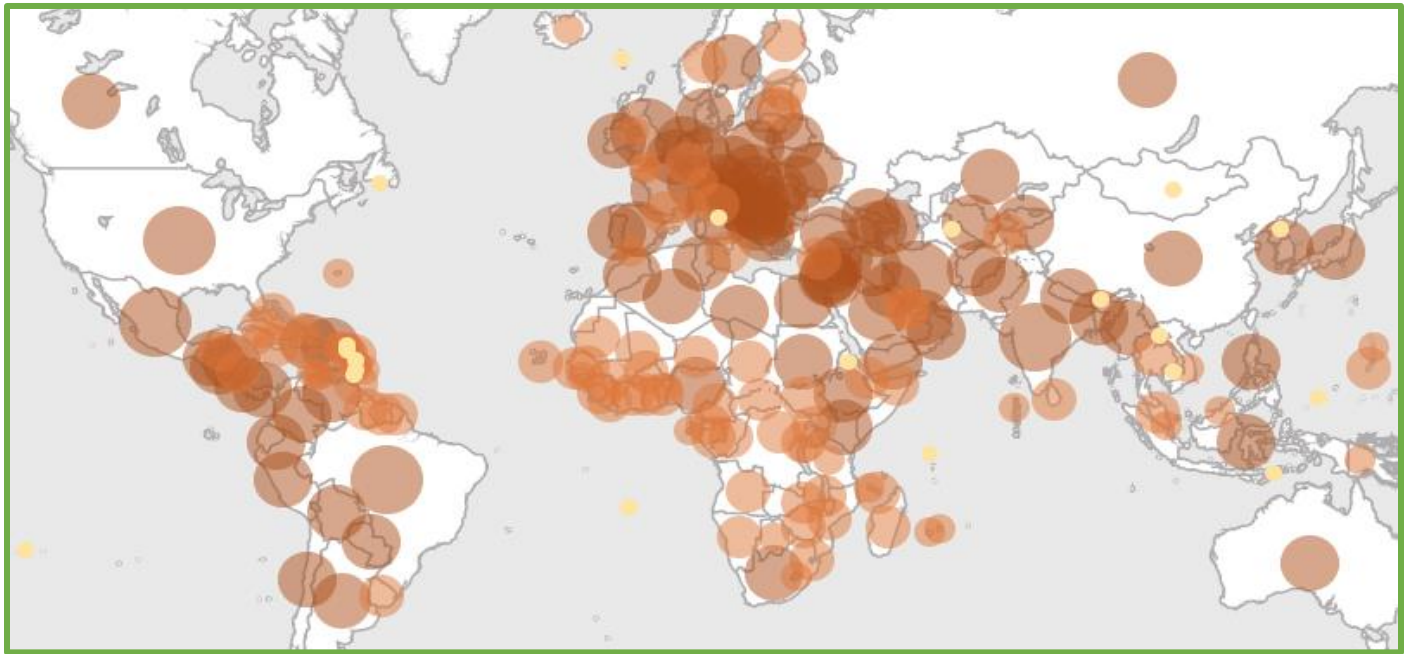
There were also some expected correlations as well, such as the less people graduated high school in the state, the more people that the state had that were living in poverty. One area I had not included in the second script I had ran as it had thrown off the curve so bad.

possible correlation:	
Population estimates, July 1, 2019, (V2019)	
Persons under 5 years, percent	
Persons 19 and over, percent	
Persons 65 years and over, percent	
Male persons, percent	
Female persons, percent	
White alone, percent	
Black or African American alone, percent	
Native american or native hawaiian / pacific islander / alask	
Asian alone, percent	
Two or More Races, percent	
Hispanic or Latino, percent	
Vetran, percent	
Persons per household	
High school graduate or higher, percent of persons age 25 ye	
Bachelor's degree or higher, percent of persons age 25 year	
With a disability, under age 65 years, percent	
Persons without health insurance, under age 65 years, perc	
In civilian labor force, total, percent of population age 16 ye	
Persons in poverty, percent	
Population per square mile - 2019	
Land area in square miles,	
FIPS Code	



possible correlation:	Virgin Islands	Vir
Population estimates, July 1, 2019, (V2019)	106,405	
Persons under 5 years, percent	7.00%	
Persons 19 and over, percent	74.60%	
Persons 65 years and over, percent	13.50%	
Male persons, percent	47.80%	
Female persons, percent	52.20%	
White alone, percent	15.60%	
Black or African American alone, percent	76.00%	
Native american or native hawaiian / pacific islander / alask	8.00%	
Asian alone, percent	0.00%	
Two or More Races, percent	2.10%	
Hispanic or Latino, percent	17.40%	
Vetran, percent	5.50%	
Persons per household	3.12	
High school graduate or higher, percent of persons age 25 ye	68.90%	
Bachelor's degree or higher, percent of persons age 25 year	19.20%	
With a disability, under age 65 years, percent	19.49%	
Persons without health insurance, under age 65 years, perc	37.51%	
In civilian labor force, total, percent of population age 16 ye	77.66%	
Persons in poverty, percent	44.23%	
Population per square mile - 2019	791.17	
Land area in square miles,	134.49	
FIPS Code	"78"	"5"

Yes, with a population that has a graduation from high school of less than 70%, and almost half of them are living in poverty. That could definitely be a concern that may want to be looked into further, but the other states that also have cases on the rise, do not have the same statistics. And if we look at the covid cases globally, we could try an ascertain that it may be a cause for other countries to have such high rates as well, however, again, there are many other locations that do not have such low graduation rates, or low poverty rates, and have the same issue as everyone else:



Realigning our attention back to the US though, as we cannot help ourselves, then there is no way that we can be of any help to anyone else.

I had dug up this bit of info, and jotted it down, but I would have needed more historical data about the cases and the growing numbers to see if there was going to be a way that I could look at as a jupyter script:

```
alabama- masking started july 16th
alaska- no masking
arizona- no masking generally, required by barbers/cosmotologists
arkansas- masking july 20th to present, children under 10 exempt
California- masking July18th to present
colorado- july 17th to present, age 10+
connecticut- april17th to present, ages 2+
delaware- april 28th to present, age 10+
WA DC- july 22nd to present, age 2+
Florida- no requirement, state policy is in conflict with major counties
georgia- no state requirement, local gov. may impose requirements as of Aug 15th
Hawaii- yes as of Nov15th statewide, state of emergency with manditory quarantine as of march 15th
Idaho- no state mandate, individual cities can and have made local requirements
Illinois- yes as of may 1st, age 2+
indiana- yes as of july 27th
Iowa- yes as of nov 17th
Kansas- yes, but largely ignored as the state legislature passed a law allowing counties to opt out of it. issued
July, reissued Nov. 18th
Kentucky- yes, july9th, ages 6+
Louisiana- yes, as of July 11th, ages 8+
Maine- original order may1st allowed for unmasking when socially distanced, revision nov4th has no exception
maryland - yes, july 31st, ages 5+
Massachusetts, partial order may2nd, revised order nov6th allows no exceptions, indoor and outdoor masking required
Michigan- yes, as of oct 5th, prior order has been invalidated by the state supreme court
Minnesota- yes as of july 25th, ages 6+
Mississippi- no, mandates have been erratically enacted and revoked for months
Missouri- no, cities have enacted local mandates in some cases
Montana- yes, as of july 15th, strengthened on nov17th
```



```
Montana- yes, as of july 15th, strengthened on nov17th
nebraska- no, required only for barbershops/solons. some cities have mandates
nevada- yes as of june24th. age10+
new hampshire- yes as of nov20th, patchwork in place prior
new jersey- yes as of july 8th, ages 3+
new mexico- yes as of july 8th, ages 3+
new york - yes as of april 17th, ages 3+
north carolina- yes, partial as of june, strengthened as of nov 23rd
north dakota- yes as of nov 14th
ohio- yes as of july 23rd, ages 10+
Oklahoma- no, despite the gov. being infected with covid himself
oregon- yes as of July 1st. ages 3+
Pennsylvania- yes, initially july1st, strengthened nov 18th
Puerto rico- yes, as of may 26th
rhode island- yes as of may 8th, ages 3+
south carolina- no, but citywide mandates are in place for many locations
south dakota- no
tennessee- no, local/city mandates are in numerous areas
texas- yes as of july 3rd, but as of dec9th 20 of the 254 have exceptions due to low case counts
Utah- yes as of nov9th, previously had a county level system based on infection rates
vermont- yes as of aug 1st. ages 3+
virginia- yes as of dec. 14th
washington state- yes as of june26th
west virginia- yes partial order since july, strengthened nov14 to remove exceptions for indoor social distancing
wisconsin- yes as of aug 1st
wyoming- yes as of dec 9th, ages 12+
```

This data was also uploaded to my GitHub. As well as all the scrubbing steps I had made with the data, and why.

But as luck happened, my bestie had come over. Her husband works at the local state prison here. They have on average about 15 employees missing every shift- and being forced into mandatory overtime. The prison is over-crowded, with 2-3 people per cell, and so social distancing is not possible.

This got me thinking. So, I opted to try and pull all the prison data I could find. There is not that good of prison data anywhere, and the data that I used the scraper tool for (which I think I did something wrong with, as after I used it, I could not open jupyter again, which is what caused the deletion, and the reinstall. And when I look at the numbers for what the prison says is supposed to be in the prison, and what is in there, the numbers are a mess. It looks like someone was just typing and putting in numbers. However, I learned that it is pretty normal for the prisoners to be hush-hush about their population. Hm, wonder why?

And when I pulled all the information about the prisons, jails, holding facilities, and went through them one by one to try and ensure that there not any duplicates, the fips codes that had the prisions, also had a positive infections rate quite higher than other areas.

And here is a break down on one of the spread sheets that I had uploaded:

I find it interesting that the states with the higher amount of correction facilities have a higher infection rate. But I still won't go as far as to make a claim that a correlation is the same as a causation. Which is why I will not have an answer to the question at the beginning. But I know that there is a population that is captive, cannot socially distance, and is a giant petri dish for potential viruses. They can potentially infect the employees, who take it home, and there is a possible explanation for the hotspots. It is getting colder, people don't go outside as much, and everyone that I see on a daily basis is sick of being isolated, and sick of masks. Which also could be another potential route that could be looked at when it comes to the infection rate of covid.

I apologize again that this is a bit late, and also that I broke anaconda. But I can promise you that I did quite a bit of back end work just to get the data to be able to do these excel files, let alone be able to present them. And now that I am at going on 53 hours with out sleep, there is no way that I am going to attempt to get jupyter back up and all the packages back in that I needed right now. Because trying to get the GeoPandas to work the first time was a nightmare. =/