

Florida COVID-19 Trends and Insights

Anna Harvey

Spring 2021

<https://github.com/anhar421/Portfolio>

Which Domain?

The domain I will be using data from is the medical field. Specifically, I will be analyzing COVID-19 data from the state of Florida.

Main site for the Florida Department of Health providing public communication, guidance, and resources. It also provides some summary data visualizations and information. The comparison of the state government's approach to COVID-19 may be interesting to compare to the data.

<https://floridahealthcovid19.gov/>

The dashboard for the Florida DOH data.

<https://open-fdoh.hub.arcgis.com/>

Data regarding overall mobility during the last year as well as some financial information. I don't know if I can access the raw data, but it may provide some insights into potential correlations between data.

<https://data.covid.umd.edu/>

Global information to use for comparison and additional insight.

<https://www.worldometers.info/coronavirus/>

The COVID Tracking Project provides compiled data from across the US. For a long period of time, they were providing more accurate information than the federal government. They also provide simplified data sources from the state and national level. Their data collection ended on March 7th but there are helpful analyses provided from the data prior to that date.

<https://covidtracking.com/>

Florida Department of Emergency Management COVID-19 page.

<https://www.floridadisaster.org/covid19/>

Johns Hopkins University site for COVID-19 information. JHU has been a consistent source of accurate data and helpful analyses throughout the pandemic.

<https://coronavirus.jhu.edu/>

Center for Disease Control COVID-19 data tracking website.

<https://covid.cdc.gov/covid-data-tracker/>

World Health Organization COVID-19 data tracking website.

<https://covid19.who.int/region/amro/country/us>

An informal, easy-to-read project to communicate common COVID-19 questions and explain data analyses. This is helpful to understand some of the more domain-specific analyses of the data. The site-

runners are primarily doctors and scientific community members.

https://linkin.bio/dear_pandemic

Which Data?

This is the primary official source for Florida DOH data. This source feeds all other data compilation sources.

<https://open-fdoh.hub.arcgis.com/>

Within this site are several COVID-19 data sets. The primary one I will be focusing on is the county data.

<https://open-fdoh.hub.arcgis.com/datasets/florida-covid19-cases-by-county>

The data shows information about cases, deaths, demographics, and testing. I could not find a codebook for the data, but I will need one. I have contacted the Florida DOH and hope to hear back from them soon. I will most likely supplement the analysis with compiled data about the state as a whole, the US, and possibly the world from one of my other data sources listed above.

Research Questions? Benefits? Why analyze these data?

My research questions are:

- Is there a correlation between population and COVID-19 cases?
- Is there a correlation between population and COVID-19 test positivity rates?
- Do rural or urban counties see more cases (adjusted for population)?
- Do rural or urban counties see more deaths (adjusted for population)?
- Are there any trends in specific counties that are different from other counties? If so, what could possibly cause those trends?

The primary focus of the analysis will be on county-by-county information. There has not been a lot of discussion of the differences between the different counties regarding COVID-19 case and death rates. My hypothesis is that certain rural counties have seen a higher overall affect from COVID-19 due to higher elderly populations and/or less masks and distancing restrictions. Understanding the affects that county populations and demographics could have on COVID-19 transmission would help justify recommendations for different types of policy changes for local governments.

What Method?

EDA will first be conducted to identify potentially important trends and outliers. Correlation measures will be conducted on variables with possible relationships. If there are outliers for case rates or other important variables in specific counties, calculations will be performed to determine if there is a statistical significance to the variations.

Potential Issues?

Florida COVID-19 data collection has been a complicated issue over the last year. There have been assertions that the Florida government has manipulated the data that has been published. This is part of the reason why I want to utilize other sources for Florida data as well as the Florida DOH. It is possible that there may be missing data that would be required for an accurate analysis. I will analyze what data is available with the assumption that it is correct to its fullest capacity.

Concluding Remarks

State-level data about the COVID-19 pandemic has been essential for state governments in their decisions about policy changes and safety restrictions. There has been little focus, however, on the differences of COVID-19 case and positivity rates between counties within the same state. Florida is a unique state with widely varying population demographics that are largely divided by county. Since the beginning of the pandemic, many of the counties have functioned under their own authorities, despite receiving different guidance from state and federal levels. The goal of this study is to investigate to what extent COVID-19 has affected Florida counties and if there is any variance between counties with different dominating demographic trends.