Andrew Zhou

Professor McDonald

Human Centered Data Science

14 December 2021

Dallas, Texas – COVID-19, Mask mandates, Unemployment, and Uninsured health insurance rates

1. Introduction

For my analysis, I will be looking at COVID-19 cases, mask mandates, unemployment rate, and uninsured health insurance rates for Dallas, Texas. Dallas is known to be a republican state and relatively against health insurance. It is common to find Dallas among the highest uninsured health insurance rates in the United States, and it has held the highest uninsured health insurance rates quite consistently over the past few years. “Hospital leaders have said that these uninsured patients put off preventable care until something is wrong and end up in the emergency room with costly uncompensated care for hospitals to deal with.” This is a serious issue on itself, however, with COVID-19 the issue may be much more severe.

It is quite common for people to get their health insurance from their employment. So, it should not be without a doubt that health insurance rates and employment rate are highly correlated values. One of the most popular topics during the period of COVID-19 was unemployment. As COVID-19 pandemic became a serious problem in the United States, the unemployment rate shot up, which likely made a lot of people lose their health insurance, and in turn these people did not seek their needed medical attention. So for my analysis, I would like to understand how COVID-19, mask mandates, unemployment rate, and uninsured health insurance rates impacted each other and ultimately how did these impact death rate?

1. Background

Before starting my analysis, I had a hypothesis of how COVID-19, mask mandates, unemployment rate, and uninsured health insurance rates impacted each other and how they impacted death rates. My hypothesis was that mask mandates and COVID-19 cases are negatively correlated. That is, when a mask mandate law is put into place, the new COVID-19 cases are likely to drop. The reason is that masks prevent water droplets that may contain the virus from spreading in to the air and prevents people from catching the virus. My hypothesis also guesses that unemployment rate and uninsured health insurance rates are positively correlated with COVID-19 cases, and that when COVID-19 cases rise, then unemployment rate and uninsured health rate will likely rise as well. I suspected this because as COVID-19 cases rise, then more businesses are likely to close down and there are less jobs open for people. Unemployment rate and uninsured health insurance rates are probably positively correlated. A lot of people have their health insurance from their employment. Also, for people who do not have insurance from their employment, they are likely to buy insurance out of pocket, but if they lose their job, they are also likely to cancel their insurance. With an even higher uninsured health insurance rate, I suspect that many people are likely to refuse seeking medical attention and increase the death rate in Dallas.

In summary, my hypothesis for my analysis is:

* COVID-19 daily cases is positively correlated with unemployment.
* Mask mandates will slow down the spread of COVID-19.
* As COVID-19 cases increases, unemployment increases.
* As unemployment increases, uninsured health insurance rates increase.
* As uninsured health insurance rates increase, more people are likely to not to get their needed medical attention.

1. Methodology

For my analysis, the main methodology methods will be using data visualization and linear regression.

1. Introduction

Why is this analysis interesting or important (to people besides you)? Does it solve a real problem or tackle an unresolved research question?

1. Background/Related Work

What other research has been done in this area? How does this research inform your hypotheses, your analysis, or your system design? What are your hypotheses or research questions?

For these COVID related questions there may not be peer-reviewed publications that are directly related to your hypothesis. There may be anecdotal claims in the popular press (blogs, newspapers) related to your analysis.

1. Methodology

Not just your analytical methods, but also, why you chose them, and how human-centered considerations such as ethics informed the way you designed your study.

1. Findings

What did you find? Use words and figures, don’t just point to code.

1. Discussion/Implications

Why are your findings important or interesting; How could future research build on this study?

This section should include a thoughtful reflection that describes the specific ways that human centered data science principles informed your decision-making in this project.

1. Limitations

This is a required section for your report.There are often many, many limitations for any study. If you honestly tried to list them all, this might end up being the longest section. You should prioritize and list the ones that are most likely to have a significant impact on your results. Specific license issues could be a limitation, depending on what data sources you used. Flaws in the data, data cleaning techniques, potential assumptions and/or how you handled missing values could be a limitation. Statistical techniques often have specific assumptions and preconditions; if you’re not certain all of the data meets those requirements - this is a good place to make that clear.

1. Conclusion

Restate your research questions/hypotheses and summarize your findings.  Explain to the reader how this study informs their understanding of human centered data science.

1. References

A list of publications (blogs, articles, research papers) that you refer to in your text.

1. Data Sources

A list of links to the relevant data sources that you used.