Presentation

|  |  |  |
| --- | --- | --- |
| Topic | Main message | Script |
| 1. Background and motivation | * COVID-19 is devasting and has wiped out 20.5 million jobs in April 2020 throughout US. * To study the impact of COVID-19 and other related aspects towards local economy. | In April 2020, 20.5million jobs in the US was removed since the pandemic. Macomb County in Michigan had an unemployment rate of 26.5% in that time. However, since vaccination was introduced in December 2020, the economic damage caused by COVID-19 has improved. In this research I will look at the relationships between COVID-19 cases, unemployment rate and vaccination rate in Macomb County. |
| 1. Problem that you’re addressing and hypothesis | To study the importance of vaccination and how that impacted the local economy, which is measured using unemployment rate as proxy. Vaccination helps with creating herd immunity, but just like seasonal flu, it comes and go. Vaccination does not directly help with | Majority of the unemployment rates impacted during COVID times were the restaurants and the services sectors, due to lockdown and reduced human traffic. My first hypothesis is that number of cases is positively correlated with unemployment rate.  Then I would like to study the relationship between vaccination, the number of cases and the unemployment rate. |
| 1. What data are you using and why |  | In this research I have used a combination of COVID-19 cases and numbers, unemployment rate and vaccination rate for Macomb County. The covid 19 cases and death data has a daily interval and is available since April 2020. The vaccination data is daily interval as well starting from December 2020 to November 2022. The unemployment rate data is updated monthly with availability from April 2020 to November 2022. |
| 1. Boring but important part: Data cleaning |  | This part takes up 80% of the data investigation part. I am using three time series of different frequency for my analysis. Unemployment rate is monthly, and vaccination and number of cases are daily. So I aggregated the latter two by month and merged all three tables together into one big table for convenience for my analysis later on. All three series have the same time range from December 2020 to November 2022 |
| 1. Semi fun part: Data exploration |  | The one on the left is the aggregated COVID-19 cases and deaths data, the middle is unemployment rate and the right is the vaccination data. There are some sort of patterns between the three. More particularly, there seems to be a closer correlation between unemployment rate and vaccination rate. COVID-19 cases and deaths do not seem to have similar patterns with the other two. |
| 1. The real deal: Data analysis |  | But we are data scientists so we want solid numbers. I have performed two tests to identify whether causality and correlation exist between the time series. Granger Causality Test in Python is used to determine whether one time series is useful for forecasting another. The null hypothesis is that the first time series does not Granger cause the second time series. Then, I proceeded with Pearson correlation test between time series. |
| 1. Results | There is no correlation between unemployment and number of cases or number of deaths. | I will be showing results for the next three slides. The table here shows the p-value from both tests. The first pair of time series I tested was number of cases and deaths with unemployment rate. They have a Pearson correlation of -0.15 but this relationship is not statistically significant in both tests. |
| 1. Results | There is no correlation between number of cases or number of deaths and vaccination | Moving on, the second relationship I tested was number of COVID-19 cases and deaths with vaccination rate. Recall that the hypothesis is they are negatively correlated. Turns out they have correlation coefficient of 0.148 and they are not statistically significant, even with a lag of 4months. |
| 1. Results | There is a strong correlation between unemployment rate and vaccination. | Ok, I finally found something! The correlation coefficient between unemployment rate and vaccination rate are 0.49 and they have p-value < 0.05! This is not as surprising as the previous two results. But essentially this means COVID-cases have no significant correlation with both unemployment and vaccination rates. |
| 1. Why do the results matter and caveat | It shows that vaccination is more effective in predicting unemployment rate than number of cases. And reason. | Anyone of you have the I am invincible feeling after your first vaccine? It turns out we are not the only one. Research done by the CDC studying the mental health of the public after getting vaccination. They feel less stressed and worried about the general situation. And that mental comfort has encouraged them to be more sociable and resume their daily lives, economy starts pulsing again and that is when we see the unemployment rate dropping. |
| 1. Human centered component and caveat | It is a study of combination of quantitative datasets and qualitative paper regarding public’s mindset of vaccination. | This research is human-centered because the findings are a combination of quantitative and qualitative research aka wide and deep type of research. The qualitative part her is the sentimental analysis and interview from CDC study. There are plenty of limitation and caveats in my research. One of them being I don’t know how many of the COVID19 cases have been vaccinated, that would have impacted my analysis. |
| 1. Conclusion |  | There is no direct correlation between vaccination or mask mandate towards number of cases and deaths. But vaccination has improved the community’s mindset to go out, economy improves, unemployment rates decreases. |

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?

COVID-19 is the first global pandemic that we have faced in recent decades. Various governments and local authorities have promptly reacted to the pandemic by exercising mandates such as masking and social distancing, which has sacrificed the economy and affected local unemployment due to dormant social activities. Its impact to overall unemployment rate is unprecedent and worthwhile to be studied for. I would like to research on the impact of COVID-19, state lockdown and implementation of mask mandate, and vaccination rate, on the unemployment rate in Macomb County, Michigan. Eventually, I would like to see the correlation between COVID-19 mandates, such as masking, and the overall economy and unemployment rate.

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 yo