

1. **Motivation/problem statement:**

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 unprecedented 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.

2. **Research questions and/or hypotheses:**

Research question: How did the unemployment rate in Macomb County change with the pandemic? Did the mask policy help with the unemployment rate? How is vaccination rate impacting unemployment rate?

Hypothesis: When COVID mandates are being exercised, unemployment rate was not impacted in the first or second week. However, as social distancing and lockdown become more prominent, the unemployment rate increases when businesses are forced to close.

3. **Data to be used:**

There are limited public datasets available about Macomb County, Michigan. [The US Bureau of Labor Statistics](#) provides the most complete monthly breakdown unemployment data. [Michigan State](#) has done a good job in collecting information about vaccination rate since June 2021([link 1](#), [line 2](#), [link 3](#)). We hope that by combining both datasets, we can create correlation between vaccination rate, mask mandate and unemployment during the period of data availability.

Limitation of dataset: The vaccination record starts from March 2021 and although it is claimed to have daily record, some records are missing.

4. **Unknowns and dependencies:**

It would be more helpful to have more comprehensive dataset around business performances during COVID-19 times. At the same time, we do not have census data about the type of employment for residents in Macomb County. Therefore, we can only make assumptions on that during that critical time when COVID-19 is prevalent, that is main factor causing unemployment. Another assumption we made while using the vaccination rate data is that businesses and daily activities start to resume normally, and that would indirectly increase business performance and demand, hence improving employment rate.

5. **Methodology:**

Data construction

The four datasets will first be downloaded, only Macomb County data will be looked at in this exercise. The three vaccination datasets will be combined into a more sequential and comprehensive list. Main tools used here will be Python Pandas and Numpy library.

Data visualization

COVID-19 cases and death, vaccination rate and unemployment rate will be assessed during the time period to uncover any related trends. The time series will be visualized with a combination of matplotlib and seaborn libraries.

Identifying correlation

Considering there will be a time lag between policy or vaccine implementation with COVID-19 cases improvement, time lagged cross correlation will be used to perform the time series analysis. [Granger causality test](#) will be used here to examine the correlation between the economic data. This Python [stasmmodels library](#) will be used to perform the analysis.

Analysis

The intended research outcome is to identify if the COVID-19 cases and policies have impacted unemployment rate and to what degree. Therefore, it is assumed that a [rolling window correlation](#) analysis should be able to show that relationship. The results will be visualized with a combination of matplotlib and seaborn libraries. A sequential correlation heatmap will be a great tool to visualize how the correlation has changed over time. And those highly correlated periods will be examined to be connected potential events such as lock down or vaccination rate milestone.

6. **Timeline to completion:**

Activity	Description	Estimated Deadline
Download, explore and construct datasets	The four datasets will be downloaded and combined to construct a more complete dataset. More specifically, the three vaccination records will be combined into a continuous time series from March 2021 to November 2022.	Nov 18th
Clean and manipulate datasets	Both datasets will be cleaned and formatted into the desired format to be used for analysis.	Nov 21st
Visualize each time series and discover any anomaly or unwanted trend	This step consists of visualizing the time series and exploring potential trend.	Nov 25th
Analyze correlation between time series	Perform Granger causality test and rolling window correlation analysis.	Nov 27th
Interpret result and prepare write up	Study the correlation map to discover explanation to the discovered trend.	Dec 4th
Prepare for presentation	Incorporate findings and analysis into presentation.	Dec 5th
Complete report and repo	Wrap up report and code to be pushed to repo.	Dec 12th