DATA 512 – Human Centered Data Science

Analysing Economic Impacts of COVID-19 in Monroe County:

Unemployment and Housing Market

Introduction

Unemployment has been a major concern since many years now. There are numerous reasons that impact unemployment – such as reduction of labor force, workforce fluctuations, technological changes, and economic conditions.
Unemployment and economy have a direct impact on each other. A high unemployment rate affects the economy in many ways. Unemployed people tend to spend less, may accrue more debt, and unemployment may lead to higher payments from state and federal governments for things like food stamps.
And when economic activity is low, firms cut jobs and unemployment rises. In that sense, unemployment is countercyclical, meaning that it rises when economic growth is low and vice versa.

Towards the end of 2019, the human race faced one of the worst pandemics of all times – COVID-19. The COVID-19 pandemic has had far-reaching economic consequences including the COVID-19 recession, the second largest global recession in recent history, decreased business in the services sector during the COVID-19 lockdowns, the 2020 stock market crash, which included the largest single-week stock market decline since the financial crisis of 2007–2008 and the impact of the COVID-19 on financial markets. The pandemic was also a factor in the 2021–2022 global energy crisis and 2022 food crises.

There have been several impacts of COVID-19 – it had widespread impact on economy in terms of GDP, unemployment, changes in housing prices, health infrastructure, food availability, amongst other impacts that are less visible – such as taking a toll on mental health of people, long term changes in work patterns as work from home gained traction, and in how people communicate with each other.

Through this project, I aim to answer questions around the economic impacts of COVID-19 in Monroe County, NY.

Given the widespread impact of the pandemic on the social and economic welfare of the human population, it seemed important to understand exactly how bad the conditions faced by the population of Monroe County was.

It focusses on a very real problem – it can help the government take actions to mitigate such situations in the future. Although controlling GDP, unemployment, and housing prices in the face of a pandemic can be impossible, certain measures can still be taken to ameliorate the situation proactively.

This is a strong human centered problem as there is a direct negative impact of unemployment on the population, and we can keep human perspective in mind while arriving at the result. I also hope to answer some important ideas that are a consequence of the pandemic that has impacted a huge chunk of the world's population in direct and indirect ways.

Background/Related Work

A number of blogs and news articles have covered the topic of how unemployment rose during COVID-19 in Monroe County. These articles point at how unemployment in April jumped to 15 percent.
On the other hand, there were certain articles that hinted at how the unemployment rate in Monroe County dropped a point, signalling continued economic recovery in the wake of the COVID-19 pandemic.

Thus, it lead me to structure the following hypothesis:

Unemployment rates in Monroe County have a positive correlation with COVID-19 cases.

There are not as many articles that talk about the median housing prices. A few articles from March 2020 talk about how coronavirus has not yet impacted the housing market in Monroe County, and that Rochester's housing market is fairly immune to the impact of the COVID-19 pandemic.

Thus, I came up with the following hypothesis for median housing prices:

Median housing prices are not "Granger-Caused" by COVID-19 cases in Monroe County, that is, they do not increase with the increase in the case count.

Data Sources

- Unemployment Rate in Monroe County, NY - This has been obtained from the U.S. Bureau of Labor Statistics and contains the percentage of people who are unemployed, as obtained from the Current Population Survey (CPS), also known as the household survey.
  The unemployment rate is the unemployed percent of the civilian labor force - [100 times (unemployed/civilian labor force)].

  [U.S. Bureau of Labor Statistics, Unemployment Rate in Monroe County, NY [NYMONR5URN], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/NYMONR5URN, November 10, 2022.]

- Housing Inventory: Median Listing Price in Monroe County, NY - This has been obtained from Realtor.com and shows the median listing price in a given market during the specified month.

  [Realtor.com, Housing Inventory: Median Listing Price in Monroe County, NY [MEDLISPRI36055], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/MEDLISPRI36055, November 10, 2022.]

- Gross Domestic Product: All Industries in Monroe County, NY – This has been obtained from the U.S. Bureau of Economic Analysis and contains the yearly GDP in units of thousands of US dollars.

  [U.S. Bureau of Economic Analysis, Gross Domestic Product: All Industries in Monroe County, NY [GDPALL36055], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/GDPALL36055, November 10, 2022.
  Data obtained from: https://www.bea.gov/data/gdp/gdp-county-metro-and-other-areas]

- COVID-19 data from John Hopkins University - The RAW_us_confirmed_cases.csv file from the Kaggle repository of John Hopkins University COVID-19 data. This data is updated daily.

For these data sources, under guidelines for citing these sources and license or terms of usage, the following has been mentioned:
Unless stated otherwise, information published on this site is in the public domain and may be used or reproduced without specific permission.

Methodology

To see if unemployment trends changed with respect to change in COVID-19 cases, I used Pearson Correlation coefficient.
The Pearson correlation method is the most common method to use for numerical variables; it assigns a value between − 1 and 1, where 0 is no correlation, 1 is total positive correlation, and − 1 is total negative correlation. This is interpreted as follows: a correlation value of 0.7 between two variables would indicate

that a significant and positive relationship exists between the two. A positive correlation signifies that if variable A goes up, then B will also go up, whereas if the value of the correlation is negative, then if A increases, B decreases.
Thus, it is an appropriate test to test if unemployment increased with increase in COVID-19 cases.

The following steps were followed:

- Cleaned the data and transformed the daily covid case count to monthly case count.
- Plotted the daily case count and the unemployment percentages with respect to time from February 2020 to October 2022 to perform a visual analysis.
- Calculated the value of Pearson coefficient between unemployment percentages and COVID-19 cases using the pearsonr() function[8] from the scipy library in Python for this analysis.

To see if the median housing prices changed with respect to change in COVID-19 cases, I plan to use Granger causality test, which is a statistical hypothesis test for determining whether one time series is useful for forecasting another. If probability value is less than any level, then the hypothesis would be rejected at that level. This test requires the time series to be stationary.

The following steps were followed:

- Cleaned the data and transformed the daily covid case count to monthly case count.
- Plotted the daily case count and the median housing prices with respect to time from February 2020 to October 2022 to perform a visual analysis.
- The stationarity of the time series was tested using Augmented Dickey Fuller Test.
- None of the time series was stationary, second order differencing is applied to both the time series to make them stationary.
- Granger causality test is applied to test if one time series (COVID-19 case count) causes another (median housing prices). grangercausalitytests() function from the statsmodels library in Python for this analysis.

Limitations

In this section, we will talk about some assumptions and limitations of Granger Causality Test. This test has the following important assumptions:

- **Linearity**

The original formulation of G-causality can only give information about linear features of signals. Extensions to nonlinear cases now exist, however these extensions can be more difficult to use in practice and their statistical properties are less well understood.

The COVID-19 case count data and the housing prices are not strictly linear functions of time.

- **Stationarity**

The application of G-causality assumes that the analysed signals are covariance stationary. Both of the time series were not stationary. Thus, we apply differencing to make them stationary before we perform the analysis.

- **Dependence on observed variables**

A general comment about all implementations of G-causality is that they depend entirely on the appropriate selection of variables. Obviously, causal factors that are not incorporated into the regression model cannot be represented in the output. Thus, G-causality should not be interpreted as directly reflecting physical causal chains.

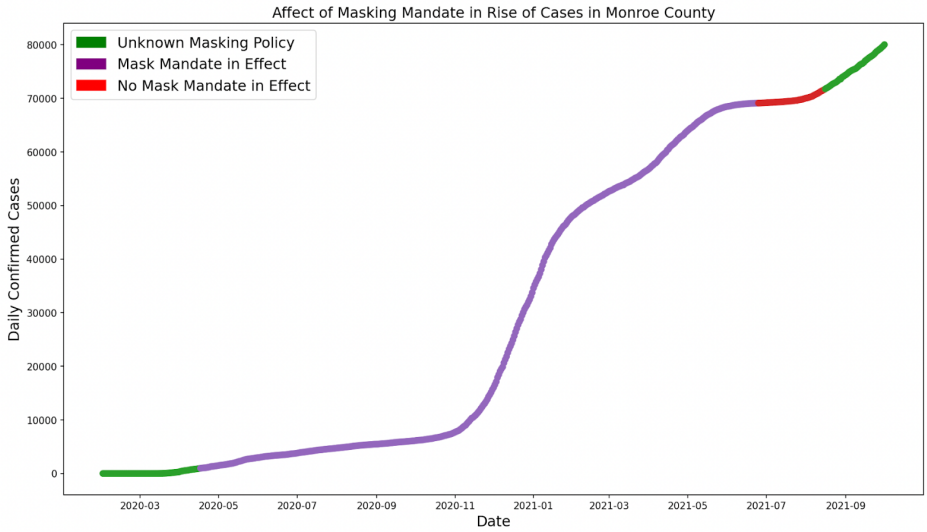
The last point holds true for our Pearson Correlation analysis as well – correlation does not imply causation! Simply because we see that unemployment percentages are negatively correlated with COVID-19 cases,

we cannot say (and SHOULD NOT SAY) that COVID-19 caused a decrease in unemployment percentages. It was a very serious and dire situation, and its impacts on economy and human population in terms of lives lost will always be looked back on with desolation.

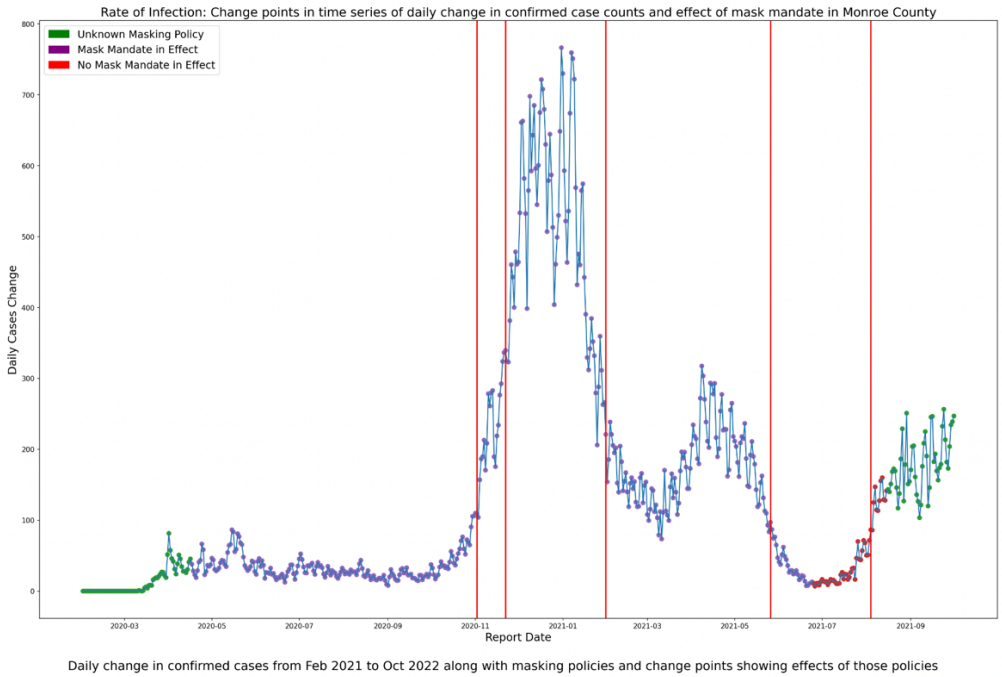
Findings

- **Visual Inspection**

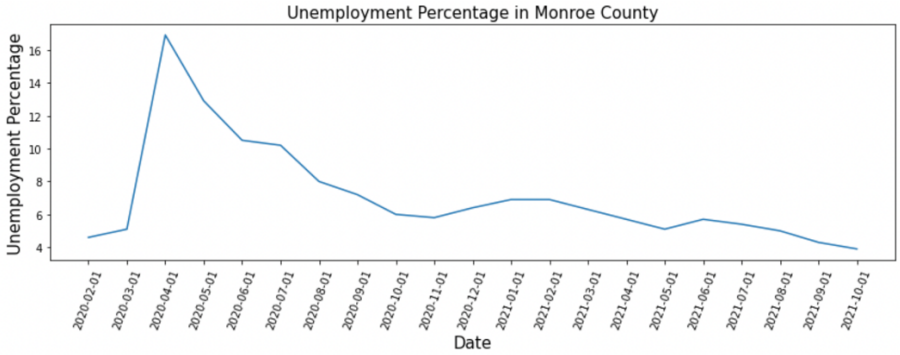
We have plots showing the trends of COVID-19 cases and unemployment percentages. As the cases rose steadily over this period, we see that unemployment percentages peaked around April 2020, which is also around the time masking policy took effect but continued to decrease as the pandemic progressed.



Plots showing trends of COVID-19 cases in Monroe County



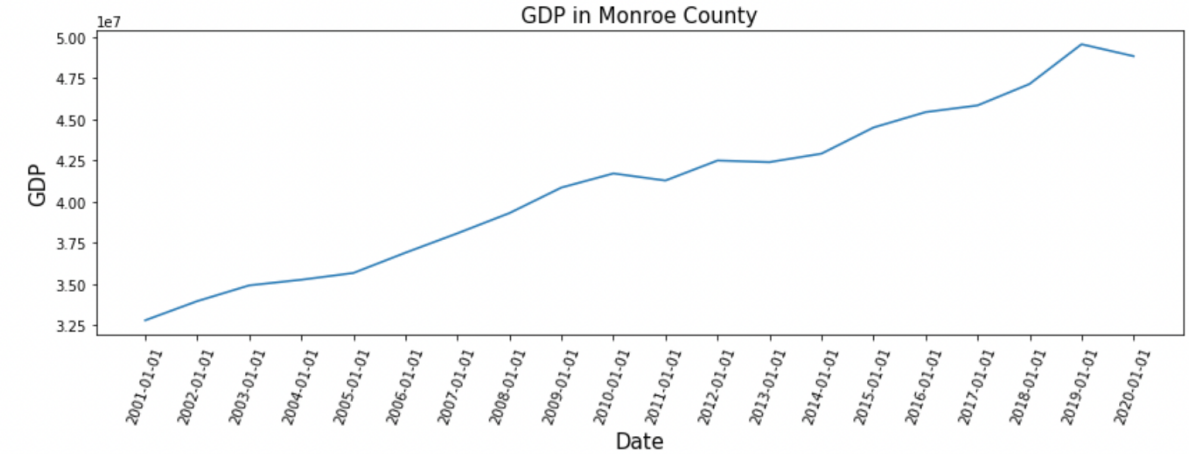
Plotting Daily New (assumed change in case count from previous day) Cases and the Mask Mandate in effect at the time along with the change points in the graph



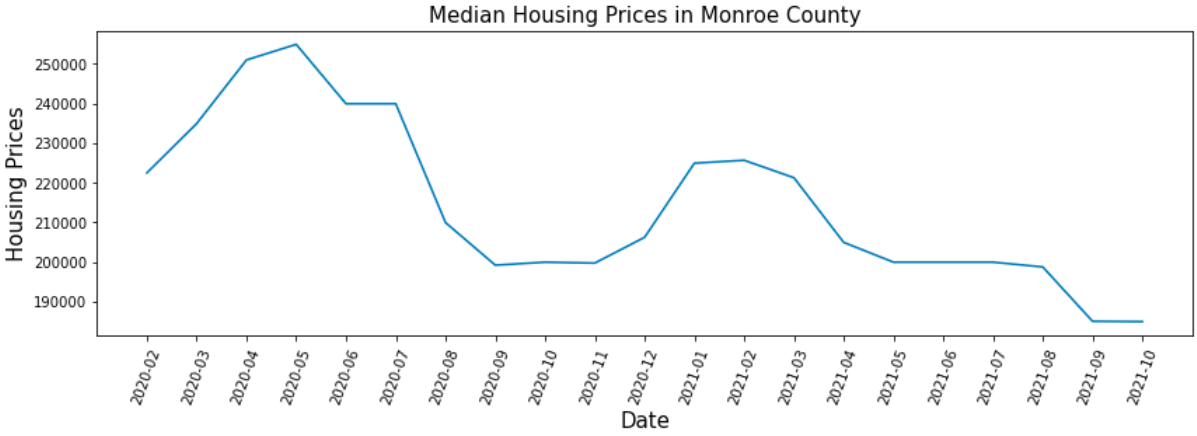
Plots showing trends of Unemployment Percentage in Monroe County

This sharp increase in unemployment percentages can be attributed to strict lockdown that was imposed in NY as a result of the pandemic, which is also when masking was mandated. Thus, unemployment rose as a result of the lockdown, but continued to recover and drop as the situation was controlled.

GDP in Monroe County mostly increased over almost 20 years from 2001 to 2019, but saw a decrease from 2019 to 2020, which is also when the pandemic was just beginning. Median housing prices had an on-off trend where they rose and fell in no discernible pattern, but after an increase at the beginning of the pandemic, they seem to stabilize and fall barring one more peak around January 2021.



Plot showing trends of GDP in Monroe County



Plot showing trends Median Housing Prices in Monroe County

- **Statistical Analysis**

Correlation between COVID-19 case count and unemployment percentages in Monroe County

The results of our statistical analysis seem to confirm the unemployment plots. We get the value of the Pearson Coefficient as **-0.533**, which signifies a negative correlation.
So while unemployment peaked as a result of the pandemic and increased case count, it decreased while the case count continued to decrease and the pandemic was controlled by new and better government policies.
Thus, the statistical analysis points at a negative correlation between COVID-19 case count and unemployment percentages.

Correlation between COVID-19 case count and median housing prices in Monroe County

We see that the p-value from Granger-Causality test is less than 0.05 for Chi-Square test (0.000) and LR test (0.0004) at lag=4, and we reject the null hypothesis that Median housing prices are not "Granger-Caused" by COVID-19 cases in Monroe County at 95% confident level.
Lag = 4 means that x(t-4) can be used to predict y(t), which in this scenario means that COVID-19 case count is correlated with and can be used to predict the median housing prices 4 months apart.
Thus, the statistical analysis concludes that median housing prices are correlated with COVID-19 case count 4 months apart at 95% confidence level.

Discussion/Implications

The findings are extremely interesting. We do not see any single pattern (strict increase or decrease) in unemployment and median housing prices. Instead, we see that they are highly susceptible to the economic climate which in turn varies with the rise and fall of the cases during the pandemic.

This is to be expected. When the cases increase drastically, it adds a load on the medical infrastructure. This domino effect continues, and has an impact on the economy and the job market. Lockdown policies that are put in place to control the spread of the pandemic renders a lot of daily wage or minimum wage workers jobless. As the condition improves and the lockdown policies are made more lenient, people step out of their houses and can resume work.

Similarly, the housing market suffers, as people put off buying houses. As a result of the economy being hit, housing prices increase, but stabilise soon enough and fall as demand for buying houses hits a low.

These findings can be used to predict the unemployment trends, as well as predict the house prices. It can be used by the agencies such as the US Labor Department and realtor companies to prepare for a situation before it occurs. While we have nearly left the coronavirus pandemic behind us, we need to be prepared for the spread of any new variant, or a new virus. Proactive measures in strengthening the medical and health infrastructure and learning from mistakes such as imposing stricter social distancing guidelines sooner can help prepare better for any employment or housing market crisis.

Conclusion

To conclude, we arrive at the following from visual inspection and statistical analysis –

- While arriving at a direct impact of masking policies on the rate of infection is difficult, it is clear to see that the masking mandate was removed after a significant drop in cases, and the cases began to rise once again after the mandate was removed. There is a lot of variation in the rate of infection – this cannot be attributed to masking policies alone. There have to be other stronger

factors at play here which include vaccinations, natural immunity after first infection, increase or decrease in testing.

- Unemployment percentages are negatively correlated with increase in COVID-19 cases in Monroe County - which was initially unexpected given the existing research and articles, but open deeper analysis, made sense because unemployment peaked once when lockdown policies took effect and stabilized over the duration of the pandemic.
- Increase in the median housing prices are Granger-Caused by the increase in COVID-19 cases in Monroe County. From visual inspection, we see that the prices rise and fall over the period of our analysis. Upon deeper inspection, we find that the housing prices first peaked when COVID-19 lockdown in New York was implemented around April 2020, and then peaked in January 2021 when we also see a sharp rise in the case count.
- Monroe County's GDP in 2020 was lower than the previous year for only the second time in 20 years starting from January 2001. This was more of an exploratory analysis and cannot be confidently attributed to COVID-19 through this project, although reports and articles suggest that the economy was hit in Monroe County during the pandemic.

The whole approach was human-centered because –

- It captures and aims to analyse the unfortunate conditions millions of people faced because of the pandemic.
- There's a direct negative impact of unemployment and rising housing prices on the population and economy.
- The techniques used in the analysis as well as the results are easy to explain to a general audience.
- The limitations of the analysis and the techniques are kept in mind instead of obfuscating the results or the numbers.
- Focus has been placed on including all necessary information around methods used and data sources for replicating and reproducing this analysis.

References

[1] https://rochesterbeacon.com/2020/05/28/jobless-rate-jumps-in-monroe-county-region/

[2] https://www.monroenews.com/story/news/2020/09/26/monroe-county-jobless-rate-drops-81-percent/42837959/

[3] https://www.democratandchronicle.com/story/marketplace/real-estate/2020/03/13/coronavirus-in-monroe-county-hasnt-impacted-rochester-ny-real-estate-market-yet/5020223002/

[4] https://nydailyrecord.com/2022/01/20/monroe-county-housing-market-shows-few-pandemic-related-warning-signs-study-shows/

[5] https://www.sciencedirect.com/topics/computer-science/pearson-correlation

[6] https://www.machinelearningplus.com/time-series/augmented-dickey-fuller-test/ s

[7] https://online.stat.psu.edu/stat462/node/188/

[8] http://www.scholarpedia.org/article/Granger_causality

[9] https://www.investopedia.com/financial-edge/0811/the-cost-of-unemployment-to-the-economy.aspx

[10] https://www.imf.org/external/pubs/ft/fandd/basics/unemploy.htm

[11] https://en.wikipedia.org/wiki/Economic_impact_of_the_COVID-19_pandemic

[12] https://builtin.com/data-science/time-series-python

[13] https://statistics.laerd.com/statistical-guides/pearson-correlation-coefficient-statistical-guide.php

[14] https://www.statology.org/granger-causality-test-in-python/

[15] https://stats.stackexchange.com/questions/133155/how-to-use-pearson-correlation-correctly-with-time-series

[16] https://docs.scipy.org/doc/scipy-0.14.0/reference/generated/scipy.stats.pearsonr.html