Social and Economic Impacts of COVID-19 on San Mateo County, CA
DATA 512 – Course Project

Nayantara Mohan
Student ID: 2129865

Table of Contents

INTRODUCTION

During the last three years, we all have been experiencing a global pandemic. This has been tragic and disruptive to many countries and has taken a deep personal toll on many individuals and their families. The datafication of the pandemic is one feature that has been difficult to ignore during the past three years. In other words, a lot of information on the pandemic's direct risk has been gathered, combined, and presented as data. With the help of this data, we have the opportunity to investigate the pandemic from a variety of prospective angles to comprehend how it has affected individuals as well as society. To be completely honest, we are just now beginning to understand and comprehend these effects.

San Francisco County, one of the 18 original counties created upon California statehood in 1850, was divided to create San Mateo County in 1856 [1]. 764,442 people were living there as of the 2020 Census [2]. As of July 2013, Wall Street Journal article identified Facebook's IPO caused a change in the US economy as San Mateo County was the home to this company and it became a top wage-earning county in the country by the end of 2012 [3]. Additionally, San Mateo County is also home to several headquarters of big companies like Visa, Sony Entertainment, YouTube, etc.

Significant inequities that affect communities of color and low-income inhabitants are present in San Mateo County. These problems are a result of complex and entrenched social and economic circumstances, but COVID-19 has made them worse. Even though there is still much to learn, it is obvious that the pandemic response must continue to focus on all low-income residents and communities of color with a specific, culturally appropriate strategy. Vaccines have been made available in San Mateo County since July 2020 as a key priority. The strategy developed by San Mateo County Health to slow the spread of the virus is intended to include personalized interventions for various at-risk community groups. The strategy is divided into two parts: 1) Improve Communication to Build Trust, Engagement, and Participation in Protection Efforts, and 2) Enhance Targeted Strategies for at-Risk Communities [4]. The steps taken by the San Mateo health officials have a human-centered approach and by this strategy, they have achieved a 95% vaccinated population (single dose) as of November 2022.

As a result of the recession brought on by the coronavirus (COVID-19) pandemic in the United States, many people left the job market, and the unemployment rate reached a peak of 13.0 percent in the second quarter of 2020. Even though significant progress was achieved against the COVID-19 epidemic by the end of 2021, the labor market had not entirely recovered. The number of unemployed declined by 4.1 million over the course of the year, to 6.8 million, and the unemployment rate decreased by 2.6 percentage points, averaging 4.2 percent in the fourth quarter of 2021, even if both indicators are still above their pre-pandemic levels. This is of concern to the people of San Mateo as well, as employment helps in basic livelihood and gives income to survive.

We hear a lot about the Covid-19 pandemic's consequences these days, including how it has hastened the move to remote work, the growth of Sun Belt tech hotspots like Miami and Austin, and the emergence of rural "Zoom towns" like Bozeman, Montana, and Truckee, California. The pandemic's contribution to America's worsening housing affordability dilemma, which has extended from tech centers and supercities

like New York, Los Angeles, and the Bay Area to cities, suburbs, and rural areas across the country, is less well known.

We have been reading about covid cases and death numbers in various news and social media platforms, but we haven't dived deep into the granular level of understanding of how this has affected the lives of people. It would be incorrect to generalize social implications at a higher abstraction as that wouldn't be fair. By doing this analysis at a county level we can understand the social and economic impact of these quantitative numbers. This is useful to understand the trend of the covid cases and deaths based on other correlated factors. E.g., the masking mandates and vaccination administered status should ideally correlate with the number of confirmed covid cases and deaths. It is practically needed to conduct this analysis to educate the masses and crowd to continue masking mandates and further encourage them to get boosters as well. Apart from that, we learn some interesting patterns of covid cases and their relation to the unemployment rate, vaccines being administered, and median prices of houses. These interesting relationships can further help in understanding on what level the lives of people in San Mateo County, CA are affected. This analysis is important for the people of San Mateo County and the government officials to help in making human-centered social and economic reforms. The insights gained in the form of data and visualizations are powerful to draw the attention of the lawmakers.

BACKGROUND

There are several works of literature talking about the effectiveness of masking to help control the transmission of COVID-19. Wearing a surgical face mask over the mouth and nose is an efficient strategy to lower the incidence of COVID-19 in community settings, according to a sizable, randomized trial conducted by scientists at Stanford Medicine and Yale University [7].

> *"We now have evidence from a randomized, controlled trial that mask promotion increases the use of face coverings and prevents the spread of COVID-19"* - Stephen Luby, MD, professor of medicine at Stanford.

This paper, along with various other literature [8],[9] promotes the usage of masks to reduce the number of confirmed covid cases and deaths as well. Thus, this forms a strong argument to further enhance the research analysis with the masking mandates being present and absent in San Mateo County. This is human centered as it is important to raise awareness of the efficacy of masking to the population to protect them from the pandemic and to minimize the death toll. A quote from the same paper [7] also states the below:

> *"Unfortunately, much of the conversation around masking in the United States is not evidence-based"* - Stephen Luby, MD, professor of medicine at Stanford.

Similar to the need of adding a masking mandate into the pandemic analysis, vaccines also prove to help in keeping the death toll and the long effects of covid in check. In a study of understanding the impact of 2

dose COVID-19 vaccines across the United States [10], it was found that vaccination reduced the overall attack rate to 4.6% from 9.0% without vaccination over 300 days.

> *"The highest relative reduction was observed among individuals aged 65 and older. Vaccination markedly reduced adverse outcomes, with non-ICU hospitalizations, ICU hospitalizations, and deaths decreasing by 63.5%, 65.6%, and 69.3%, respectively." [10]*

The conclusion of this paper [10] was that vaccination can have a substantial impact on mitigating COVID-19 outbreaks even if there is limited protection against the infection i.e., lack of masking mandate.
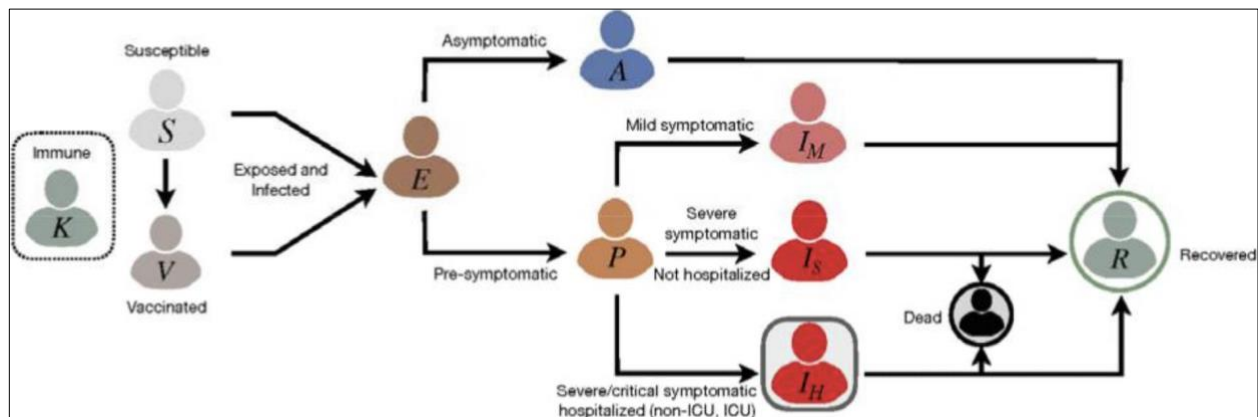

Figure 1: Schematic model diagram for infection dynamics and natural history of disease [10]

From a study published in nature [11] on the impact of vaccination on the COVID-19 pandemic in US states the below analysis was revealed which is quite interesting. The red line indicates the fraction of the population infected over time and the solid blue curve is the cumulative vaccination coverage in the population with at least one dose of vaccine. The dashed blue curve is the cumulative vaccination coverage of fully vaccinated individuals. It can be clearly observed that the fraction of infection over time reduces beyond the blue lines.
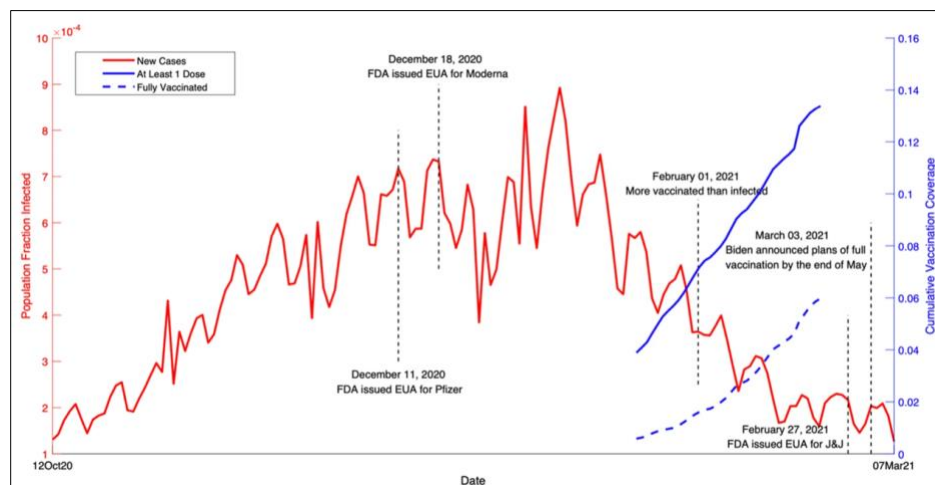

Figure 2: Impact of vaccination on the COVID-19 pandemic in U.S. states [11]

The above sets a good background to the research in this domain and it helps in informing and creating certain human-centered research questions and hypotheses for the analysis of the COVID-19 pandemic in San Mateo County. The above research has been for a subset of people throughout the United States. To make this research more granular and specific to San Mateo, we raise the following questions which directly or indirectly have a socio-economic impact on the people.

I have asked the following research questions and hypothesis for the analysis of this project: -
1. How did masking policies change the progression of confirmed COVID-19 cases from February 1, 2020, through October 1, 2021?
2. What other trends can the masking mandate by San Mateo County reveal about the pandemic?
3. With every stage of vaccines administered, is there a decrease in the number of deaths caused by COVID?
4. How was the housing rate affected in San Mateo County due to the pandemic?
5. How was the unemployment rate affected in San Mateo County due to the pandemic?

These five questions asked above are human-centered and of utmost concern to the well-being of the people of San Mateo. Based on the above background research, it is evident that masking and vaccines are for the social good of people and it does have a positive impact. But it is also true that people in San Mateo County are skeptical about masks and vaccines thus there is a need for conducting a county-level analysis to understand the level of masking adoptions and vaccines being administered. Coming to the economic impacts of the pandemic based on the unemployment levels and housing prices, it would be a good analysis to see if the two are in favor of the people or not.

These quantitative and qualitative analysis matter as these are indicators of the social and economic state of the people in San Mateo County. By leveraging these indicators, government officials can make/change policies that have favorable outcomes in the interest of the masses.

METHODOLOGY

Data Acquisition
The first step to conduct the analysis for the above-stated research questions was to find the appropriate data sets which are non-copyrighted and open to use. It is also important to note the limitations of these data sets (described further in the limitations section of the report). To illustrate a story about the pandemic, the two important datasets needed for this were the number of confirmed cases and the number of deaths. Although this data need not be accurately represented, we assume it is and conduct the analysis by adding the masking mandate, vaccines administered, housing rate, and unemployment rate data onto it.

Exploratory Data Analysis
Before I can deploy any statistical or visual analysis, it is crucial to ensure the dataset assumptions are set out and based on that the data has been cleaned and is ready to use. In this project analysis, I have used Python 3, namely the pandas, scipy, and matplotlib [14] libraries extensively. Looking at the summary statistics of the datasets and accounting for any abnormality (discussed further in limitations).

I have also converted the covid daily cases into weekly data as the trend line would be less noisy and thus easier for users to understand the trend and seasonality better. No intermediate data files were created.

Change Point Detection Algorithm

For a non-tech audience to understand a time series trend plot, it is crucial to break down the time series trend based on when a change is happening. Even though it is visually clear in some instances, but that is not always true. Hence a change point detection algorithm [12] is deployed to detect abrupt changes in the time series data to analyze the reasoning behind it. These change points are plotted as vertical lines in a distinct color for visual ease and understanding from an end-user perspective.
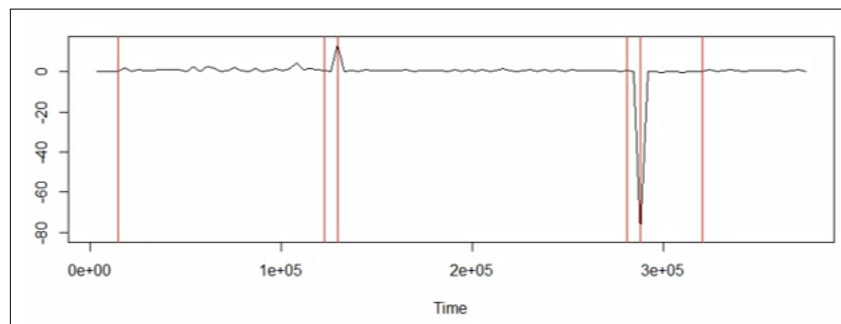

Figure 3: Graphic generated using Pelt Algorithm (Changepoint detected in red lines)[15]

I plan to use this algorithm on the weekly new covid cases data and on the daily deaths data. These vertical lines and their spacing of them will help the users understand when and how abruptly the rate of new cases or deaths is changing. Treating these change points as phases and overlaying the masking or vaccine data can help in further answering why there was a plausible change point in that given time frame.

Statistical Analysis for Hypothesis Testing

Testing the hypothesis or rather answering the question of if the unemployment rate and housing rates effect due to the pandemic can be done by setting an appropriate null hypothesis and using the student's t-test method like Welch's t-test in this case [13]. For this research purpose, we have two null hypotheses.

H_{o1} = There is no relationship between unemployment rates before and after the pandemic.
H_{o2} = There is no relationship between housing rates before and after the pandemic.

To reject the above two null hypotheses, we need to get a p-value of less than 0.05. Before we conduct Welch's t-test (which can be done on datasets with unequal variances), we gather two groups of data. We take 18 months before the pandemic and 18 months through the pandemic as the two groups of unemployment and housing rate data. Thus, on these two groups, the t-test is conducted.

This test is human centered as the result of this test gives a direct answer if more intervention is required on understanding what the government could do better to help the people with affordable living or help in increasing more employment opportunities.

Data Visualization

> *"The greatest value of a picture is when it forces us to notice what we never expected to see."* –
> John Tukey

Quoting John Tukey, it is true that the power of data visualization is very strong, and it helps in bringing out some key nuances and helps in telling a story to the audience. All the analysis will be visualized and inferred based on the story told by the data visuals. It is easy for the users to comprehend as well instead of showing code outputs or tables.

Documentation for Reproducibility

To ensure human-centered data science is practiced, the Jupyter notebook is well documented with comments and instructions. It also has relative address paths and intuitive variable names to ensure the code is reproducible and replicable. Links to all the data sources and literature surveys are also enlisted clearly. To help in facilitating further research, it is essential to make reproducibility easier. To avoid any copyright issues, the licenses, and terms of use of each dataset are also stated in this documentation.

FINDINGS

Question 1: How did masking policies change the progression of confirmed COVID-19 cases from February 1, 2020, through October 1, 2021?
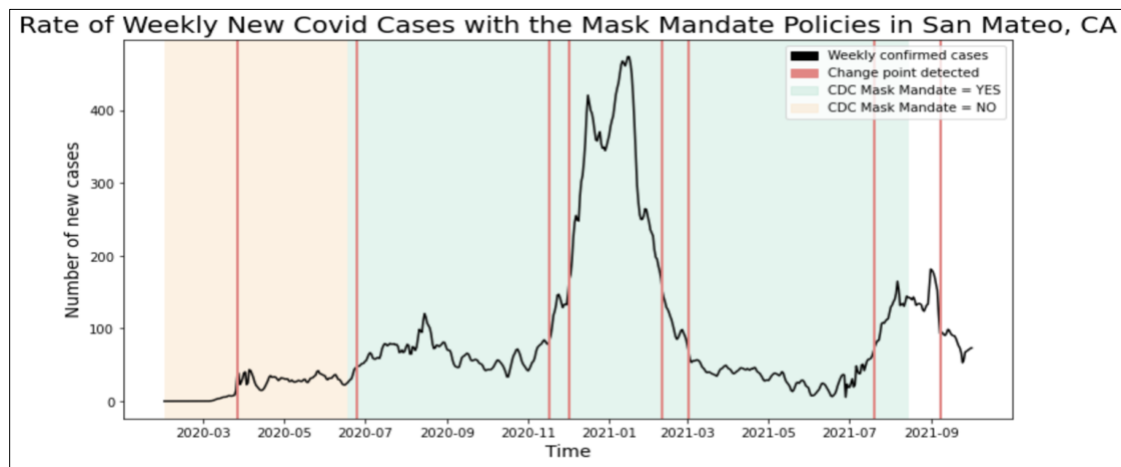

Figure 4: Effect of confirmed covid-19 cases based on the CDC masking mandates

Observation: I have used a color-blind-friendly palette such that the graphs are inclusive, and the colors can be distinctly seen by everyone. In the above graph, we can see that the CDC masking mandate has been there from late June 2020 up until Aug 2021. There are two significant peaks, namely one between Dec 2020 – Mar 2021 and another from Jul 2021 – Sep 2021 (as denoted by the red vertical lines). On further analysis by getting the San Mateo County masking mandate information and superimposing that over the CDC masking mandate, we get an interesting observation as seen in Figure 5.
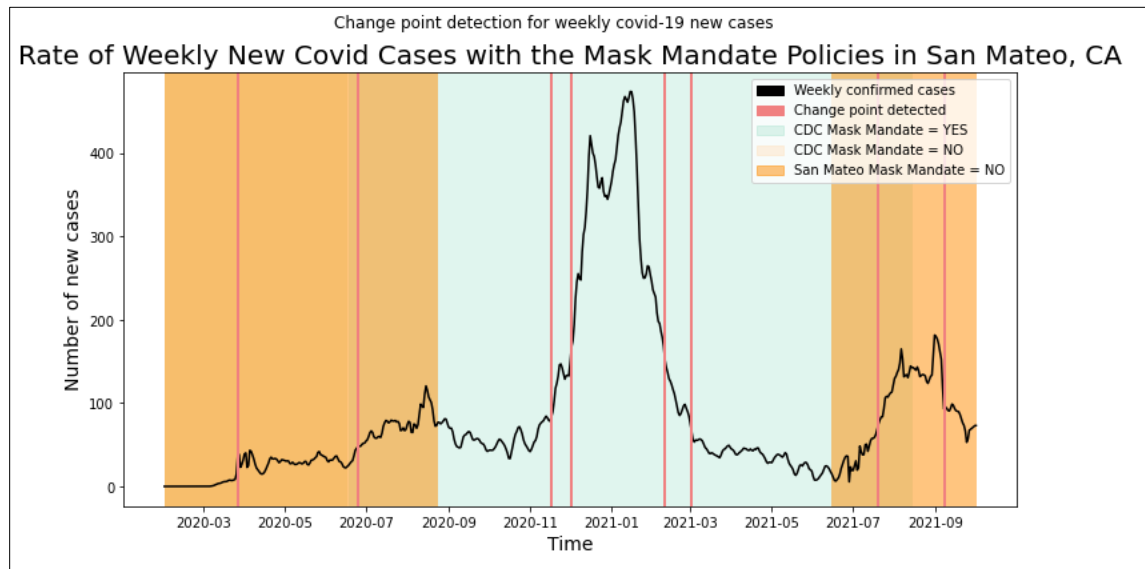
Figure 5: Effect of confirmed covid-19 cases based on the CDC and San Mateo County official masking mandates

Observation: A clear observation that can be seen is that even though the CDC masking mandate was there, the officials of San Mateo removed the masking mandate [15] in late June 2021, and thus it can be seen that the new wave after that had started cause of that. While CDC masking mandate data wasn't sufficient to explain this, on further analysis of qualitative data we can justify this spiking trend.

Question 2: What other trends can the masking mandate by San Mateo County reveal about the pandemic?
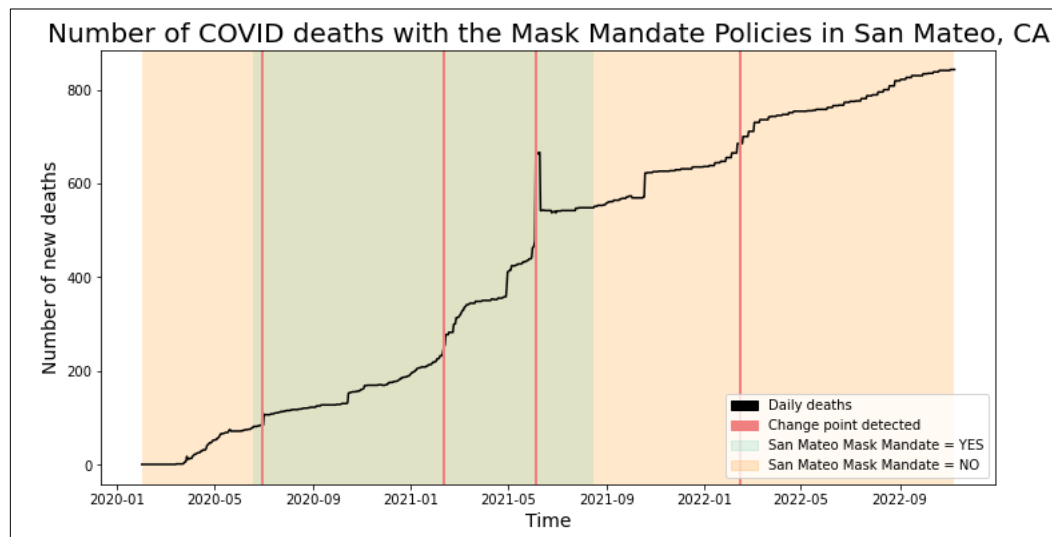

Figure 6: Effect of the daily death cases and masking mandates

Observation: I wanted to see if masking mandates had any effect on the number of deaths. Unfortunately, there is no such story being told here. As we can see that the death rates are increasing abruptly (three change points) during the phase of masking mandate being there and reduces after the masking mandate

is uplifted. It is strange how the rate of increase in deaths started to reduce. There must be some other metric that is responsible for this trend which isn't just the masking mandate.

Question 3: With every stage of vaccines administered, is there a decrease in the number of deaths caused by COVID?
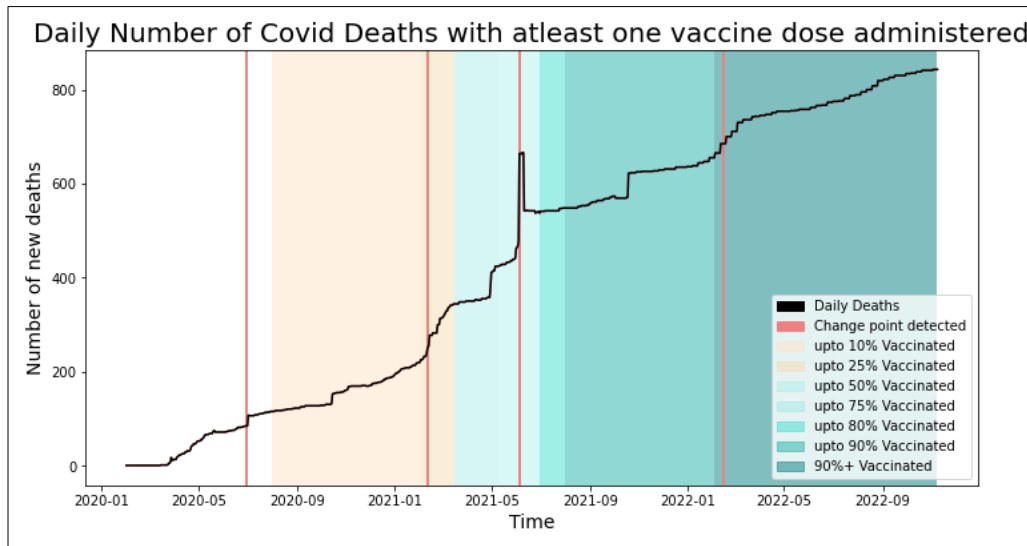

Figure 7: Effect of the daily death cases based on the percentage of vaccines administered

Observation: On taking the percentage of vaccine-administered data from the San Mateo official website [16], and plotting the same with the number of death cases, it reveals a beautiful story that falls in line with our background research about vaccines being effective in controlling the death rates. It can be seen that the abrupt changes and rise in death toll occur up until 50% of the population is vaccinated (three change points detected) and after that, there is only one change point detected. That means that the effect of the vaccine has a significant message to tell the masses.

Question 4: How was the housing rate affected in San Mateo County due to the pandemic?

Statistical test:

To conduct Welch's t-test for this group of data, we assume that the data is normally distributed (This is a strong assumption and has its limitations). As explained in the methodology, I have taken 18 months prior to pandemic data in group 1 and 18 months in the pandemic data of the housing rate data in group 2. On conducting a t-test on these two groups, we get the following metrics: -

p- value	0.75
t-value	0.317

From the above table, it can be seen that since the p-value is > 0.05, the results are not statistically significant hence we do not have enough evidence to reject the null hypothesis. This means that the housing rate is not affected by the pandemic i.e., it is random and the spikes we see it following the new cases data are by absolute chance.
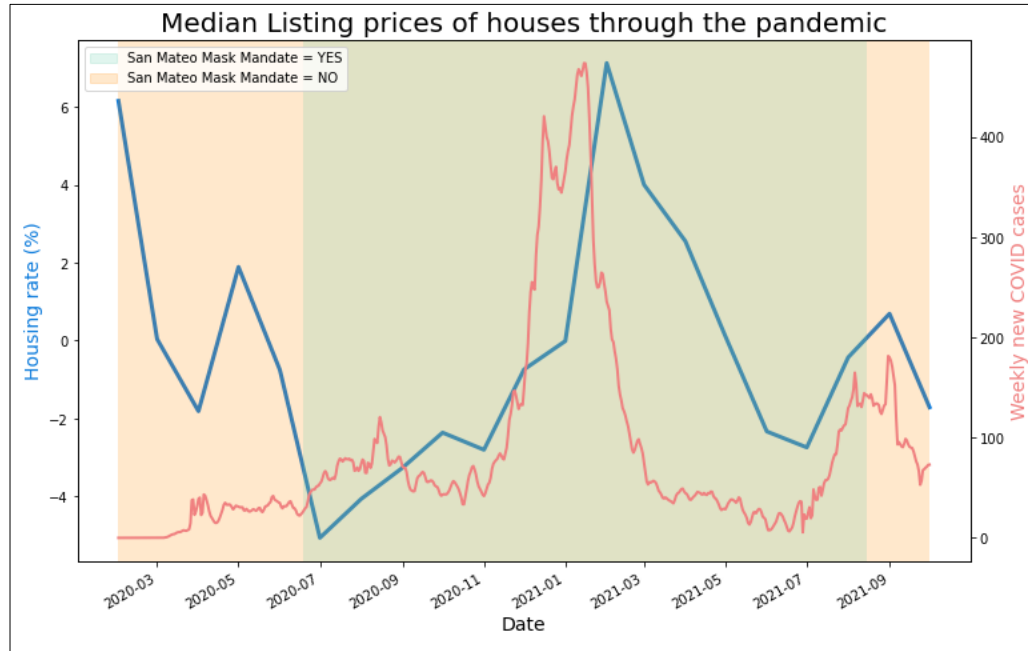

Figure 8: Effect of housing rate during the pandemic with weekly new covid cases

Observation: By looking at just the visual plot of the housing rate (blue) and new COVID-19 cases (red), it can be seen that there isn't any significant rise or fall happening due to the pandemic. There are ups and downs resembling the peaks of the waves but that's random and there are peaks even before the first wave. Hence, it's unsure what affects the housing rate.

Question 5: How was the unemployment rate affected in San Mateo County due to the pandemic?

Statistical test:

To conduct Welch's t-test for this group of data, we assume that the data is normally distributed (This is a strong assumption and has its limitations). As explained in the methodology, I have taken 18 months prior to pandemic data in group 1 and 18 months in the pandemic data of the unemployment rate data in group 2. On conducting a t-test on these two groups, we get the following metrics: -

p-value	0.00000067
t-value	-7.082

From the above table, it can be seen that since the p-value is << 0.05, the results are statistically significant and hence have enough evidence to reject the null hypothesis. This means that the unemployment rate is affected by the pandemic.
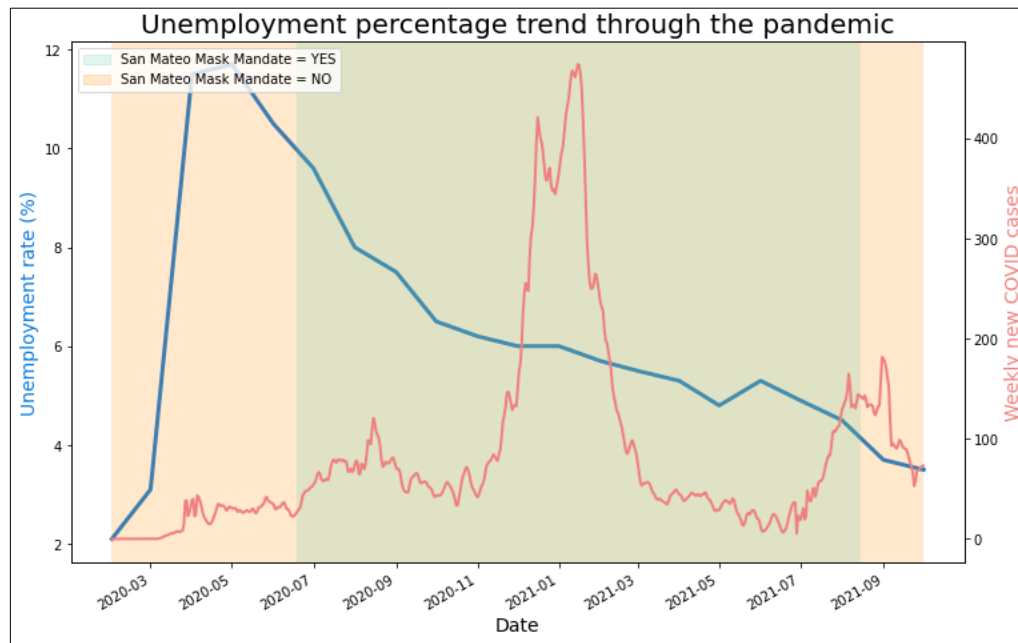

Figure 9: Effect of unemployment rate during the pandemic with weekly new covid cases

Observation: By looking at just the visual plot of the unemployment rate (blue) and new COVID-19 cases (red), it can be seen that there is a significant rise during the start of the pandemic and eventually it drops as the pandemic progresses. Hence the pandemic has influenced the unemployment rates.

DISCUSSIONS

The findings are quite important and have a human-centric value attached to them as well. Starting from the findings of new covid cases reduced based on the masking mandate, this is in line with the background research and the papers [7][8][9] that proved that masking help in controlling the transmission of new covid cases. The finding is quite interesting as it can be seen that the second wave in the covid cases after May 2021 can be attributed to the fact that San Mateo County lifted the masking mandates. Previously with just the CDC masking mandate, this trend isn't visible. This is crucial for the officials in San Mateo to understand and make human-centered decisions and impose masking mandates as and when needed for the betterment of the people. Based on the background research, it is also proven that masking is especially important for people aged above 65.

The administration of just a single dose of vaccines to the people of San Mateo County also reveals an interesting trend when the number of death cases has started to remain constant and not rise abruptly like before. NOTE: for this analysis, the data has been taken up to Nov 2022 (unlike Oct 2021 for other analysis), this is done so that we can observe the effect of the vaccines being administered well by adding in more

data points. Given this analysis only takes into account that the people have been administered just one dose, it is crucial to see and note the effect of being fully vaccinated and boosted as well.

The housing rate and affordability of basic living are essential to the people of San Mateo County. The rates have been very abrupt and noisy. There is no correlation with the pandemic as it has been unpredictable even before the pandemic. Further analysis and enlisting other possible economic factors that could be responsible for this trend are needed. This is crucial from a human-centered perspective to ensure correct economic reforms are being taken by the officials. The same goes for unemployment, it is good to see that the rate has been dipping and this doesn't seem to pose a concern for immediate action to be taken.

Further research can be built on all these analyses. All the code and documentation have been presented with keeping the ease of reproducibility in mind. In the case of masking, enhanced research on the use of masking on elderly people will be useful to understand the implications of masking on the death rate. The vaccine analysis in this project is limited to just a single dose being administered. Further analysis of fully vaccinated and boosted individuals can give a stronger correlation and evidence of vaccination. This is a human-centered way of presenting factual data to the masses to spread the importance of vaccines. It can also help anti-vax individuals to consider it.

LIMITATIONS

There are several limitations to this study in terms of the assumptions made and the collection of data. The below section details most of the limitations encountered while performing this analysis.

1. Data collection error in confirmed covid cases: In the RAW_us_confirmed_cases.csv dataset, it was observed that cumulative numbers of the total number of confirmed cases have been reported. While filtering the same for a county level i.e., San Mateo, it was observed that when the cumulative frequency was converted to a daily count, there were few instances of negative values. This isn't possible if the cumulative count was done accurately. However, this was the case for only 0.05% of the data, which is a small fraction thus it was ignored, and these data rows were deleted. To handle this, a weekly average was also taken to ensure such abnormalities do not affect the authenticity of the dataset.

2. NaN values in masking mandate: The rows having NaN values for masking mandate were replaced with "No" as the dataset had only "Yes" and "NaN" values and the dataset description did mention it being a "Yes" or No". I also cross-validated if the CDC masking mandate was a "No" by reading about it further online (Note: this was for the time period before masking mandate came into the picture, i.e., during the start of the pandemic) [15]. Thus, it seems accurate to make this assumption, but it is a limitation of the dataset.

3. CDC masking mandate only until Sept 2021: Another limitation of the dataset is that the CDC masking mandate is only there until Sept 2021. This means that we cannot assume if the masking mandate existed or not after that time. To help ease our analysis, I leveraged qualitative data found

in announcements and noted the masking mandates kept by the San Mateo officials to fill in this gap in the data [15].

4. <u>Limitation of The New York Times Survey Data:</u> Although the data provided by the New York Times survey data is very essential and has the required elements of human-centeredness as it gives information on how often people wore masks in a particular county, it lacks data. The data is from just one week which is of no use and no significant analysis could be done from it.

5. <u>Accuracy of COVID deaths data:</u> There is a lot of background research about the deaths being classified as COVID-19-related dates is questionable. It is quite ambiguous to relate death to one cause, especially for individuals with comorbid conditions. Thus, a clear assumption that needs to be made while using this dataset is that it is assumed that the cause of death and the death toll is because of COVID.

6. <u>Assumption of Masking:</u> Despite the data showing that a masking mandate was there or not there, it has to be assumed that 100% of the population was masked during the masking mandate "Yes" and 0% of the population was masked during the masking mandate "No". This isn't a fair assumption to make as in the real world the percentages could vary drastically, and it is solely based on an individual's choice to be masked even if there isn't a mandate.

7. <u>Side effects of Vaccines:</u> Although the efficacy of the vaccine has proven to be very high, there is a small section of the population (immune compromised) who may have reactions that are not favorable to taking a vaccine. For such populations, a different course of action needs to be devised to help in keeping them safe during the pandemic outbreaks.

8. <u>Assumption of data normality:</u> As stated in the findings, Welch's t-test can be performed only under the assumption of a data normality meeting. In the case of housing rate data, only a part of the data is normal and in the unemployment rate, the data is skewed normal which are not the ideal conditions to perform a t-test. However, since other assumptions to perform the t-test were met, it was performed. However, a note has been mentioned to not take the p-value as the result and to rely on the simple data visualization which delivers the story more accurately.

9. <u>Limited information on variants in the study:</u> We know that every variant of the pandemic has been of a different kind in terms of infection and transmission rates. It would be incorrect to assume the entire study is based on a single variant and the same effects are seen. Thus, future analysis should account for the kind of variant and see the wave accordingly. For e.g., the omicron variant hasn't been accounted for in this research.

10. <u>Equity in vaccine distribution:</u> A strong assumption that has been made is that every individual has equal access to getting a vaccine. This is a concern as in reality it probably is not true and it is an unethical form of practice to favor a certain population more than the other while providing vaccines.

CONCLUSION

The study reveals interesting results from the analyses conducted on the following research questions and hypotheses.

How did masking policies change the progression of confirmed COVID-19 cases from February 1, 2020, through October 1, 2021?

>This study was conducted by not just including the CDC masking mandate data but also going one level further and understanding the San Mateo County masking restrictions by reading up on their official website. This human-centered approach reveals the trend of the second wave being present due to the San Mateo government officials easing and lifting their masking update. The background research and study together do recommend continuing to wear a mask when there are future outbreaks or for any other airborne pandemic in the future.

What other trends can the masking mandate by San Mateo County reveal about the pandemic?

>This part of the study reveals that just masking isn't helpful in staying safe from the pandemic, it is essential to look for other possible indicators that are leading up the death rates. One such indicator is being vaccinated as this has a correlation with reducing the death rates. However, there is scope for further human-centered analysis to be done to gain insights into other factors.

With every stage of vaccines administered, is there a decrease in the number of deaths?

>One of the most convincing analysis outcomes that can be seen in this part of the study was that the death toll does rise much slower once people are vaccinated with at least one dose of vaccine. This is a crucial result as vaccination is very crucial for individuals and this information needs to be spread throughout the masses.

How were the housing rate and unemployment rates affected in San Mateo County due to the pandemic?

>The conclusion from this analysis was that the housing rate isn't affected by the pandemic whereas the unemployment rates have been impacted. However, after a sharp rise, it has only been decreasing which is a good sign for the people of San Mateo.

All these analyses and answering the above questions, help a reader inform the understanding the human-centered approach taken for all the analyses. The need for masking mandates and vaccines helps in the social impact and the outcome of the housing rate and unemployment rate helps in understanding the economic impact. All this information that has been analyzed through quantitative and qualitative data can help San Mateo County officials implement the appropriate reforms for the betterment of the people.

REFERENCES

[1] "San Mateo County". *Geographic Names Information System*. United States Geological Survey, United States Department of the Interior. Retrieved February 1, 2015.

[2] San Mateo County, California". *United States Census Bureau*. Retrieved January 30, 2022.

[3] Wikipedia contributors. (2022, November 27). San Mateo County, California. In *Wikipedia, The Free Encyclopedia*. Retrieved 23:59, December 11, 2022, from https://en.wikipedia.org/w/index.php?title=San_Mateo_County,_California&oldid=1124051698

[4] San Mateo County Health, https://www.smchealth.org/covid-19-vaccination-program-overview

[5] "Effects of COVID-19 pandemic on the Employment Situation news release and data" (U.S. Bureau of Labor Statistics, last modified January 7, 2022), https://www.bls.gov/covid19/effects-of-covid-19-pandemic-and-response-on-the-employment-situation-news-release.htm.

[6] "How the rise of the rest became the rise of the rents", https://www.bloomberg.com/news/features/2022-09-08/why-did-housing-costs-explode-during-the-pandemic

[7] *"Surgical masks reduce COVID-19 spread, large-scale study shows"*, Krista Conger, Stanford Medicine News Center, https://med.stanford.edu/news/all-news/2021/09/surgical-masks-covid-19.html

[8] An evidence review of face masks against COVID-19, Jeremy Howard, Austin Huang, Zhiyuan Li, et al.

[9] Echu Liu, Samantha Arledge, Individual characteristics and demographics associated with mask wearing during the COVID-19 pandemic in the United States, Asian Journal of Social Health and Behavior, 5, 1, (3), (2022).

[10] Moghadas, Seyed M et al. "The impact of vaccination on COVID-19 outbreaks in the United States." *medRxiv : the preprint server for health sciences* 2020.11.27.20240051. 2 Jan. 2021, doi:10.1101/2020.11.27.20240051. Preprint.

[11] Chen, X., Huang, H., Ju, J. *et al.* Impact of vaccination on the COVID-19 pandemic in U.S. states. *Sci Rep* 12, 1554 (2022). https://doi.org/10.1038/s41598-022-05498-z

[12] *"Detecting the Change Points in a Time Series"*, Chris Kuo, https://medium.com/dataman-in-ai/finding-the-change-points-in-a-time-series-95a308207012

[13] *"Welch's t-test"*, Python for Data Science, https://pythonfordatascienceorg.wordpress.com/welch-t-test-python-pandas/

[14] *"Matplotlib Cheat sheet"*, Matplotlib Org, https://matplotlib.org/cheatsheets/_images/cheatsheets-1.png

[15] San Mateo Health, County reopening, https://www.smcgov.org/san-mateo-county-reopening

[16] *"Vaccination totals and locations data"*, San Mateo Health, https://www.smchealth.org/post/vaccination-totals-locations-data

DATA SOURCES

1. COVID-19 confirmed cases from John Hopkins University ([Link](#))
License: [Attribution 4.0 International (CC BY 4.0)](#)
Description: Total number of daily confirmed covid cases across various counties in the United States.

2. COVID-19 US deaths from John Hopkins University ([Link](#))
License: [Attribution 4.0 International (CC BY 4.0)](#)
Description: Number of daily deaths in the United States based at a county level.

3. CDC masking mandates by county ([Link](#))
Terms of Use: [Use of Agency Materials](#)
Description: The CDC masking mandate at a county level was collected until September 2021.

4. The New York Times mask compliance survey data ([Link](#))
License: [free-of-cost, perpetual, non-exclusive license](#)
Description: Estimate of the mask usage by counties in the US based on roughly 250K interviews conducted by Dynata from July 2 to July 14, 2020.

5. Housing median price month-over-month ([Link](#))
Terms of Use: [FRED® Services Terms of Use](#)
Description: This dataset contains the median listing price in a given market during the specified month for San Mateo County.

6. Unemployment rate in San Mateo ([Link](#))
Terms of Use: [FRED® Services Terms of Use](#)
Description: The units of unemployment are in percent and the frequency is monthly. The unemployment rate is the unemployed percent of the civilian labor force [100 times (unemployed/civilian labor force)].