The Mask Mandate Efficacy for Containing COVID-19 within Michigan

Submitted by

David Piwowarski

Applied Statistics

To

The Honors College

Oakland University

In partial fulfillment of the

requirement to graduate from

The Honors College

Mentor: Hon Yiu So, Assistant Professor

Department of Mathematics and Statistics

Oakland University

April 7th, 2023

Abstract

This project extended previous research on face mask policy and the incidence of COVID-19 to be applied to the population of Michigan. With the ongoing coronavirus pandemic, mask mandates seemed to be a necessary containment solution that could suppress COVID-19 cases and reduce pressure on health care services. To estimate the efficacy of the mask mandate in Michigan, we analyzed the COVID-19 datasets from the Centers for Disease Control and Prevention (CDC), focusing on Michigan's counties. Furthermore, we also considered additional data sources to enrich the surrounding discussion regarding mask efficacy, namely how the rest of the United States fared through the pandemic. We applied the Generalized Linear Mixed Models (GLMM) to compare counties when masks were mandatory in public and when masks were compulsory. We also provided recommendations to public health officials over the state of the pandemic within Michigan's locale.

1. Introduction

Masks have been part of a contentious debate within the medical world and now within societal discourse itself over their effectiveness in preventing viral infections. Over the past two decades, research on the prominent appearance of contagions, such as H1N1 influenza, commonly referred to as swine flu, SARS (Severe Acute Respiratory Syndrome), and MERS (Middle-East Respiratory Syndrome) has amplified pandemic concerns and the discussion over whether masks should be worn to protect the uninfected populace from further spread. Specifically, the conversation has veered into the territory of whether healthy individuals should be mandated to wear masks regularly to reduce general spread and the associated deaths that come with it. It is reasonably understood that vaccines and other pathogen-targeted medications won't be available at the beginning of a pandemic event to tackle the illness.

Recently, COVID-19 (SARS-CoV-2), informally derived from coronavirus, has become a monumental global health crisis for the public health community. Coronavirus is primarily transmitted through direct contact, where the virus travels person-to-person by coughing and sneezing, and indirect contact, to a lesser extent, where contaminated surfaces are touched and not disinfected as time progresses. In addition, the virus has an element of being asymptomatic—individuals without any obvious signs or symptoms who carry the virus unknowingly—that can further broach unsuspecting subjects who do not recognize seemingly healthy individuals as vectors to spread coronavirus. Finally, the lack of medical supplies posed a challenge in the initial treatment of patients at the beginning of the health emergency. In particular, a lack of abidance to early warnings and inaction by governments and health organizations to curb the spread in its infancy have made it highly lethal (Aledort et al., 2007). Since the beginning stages of the worldwide pandemic in March 2020, government officials have rushed to construct efficacious policies to contain the spread of this dangerous virus without the time or preparations made in advance.

With the emergence of various SARS-CoV-2 variants, the worldwide infection rate has skyrocketed. With only non-pharmaceutical methods being left at their disposal, FTTIS (Find, Test, Trace, Isolate, and Support), otherwise known as the "Zero-COVID" policy, was implemented in a variety of countries, including the United States, with the intent of preventing further viral transmissions upon first detection of COVID within an identified locale. These actions included lockdowns to stunt population movement, contact tracing to track individuals who may have encountered an infected individual or come into close proximity of a highly infected area, and repeated testing to isolate further cases (Chung et al., 2021). However, with prolonged efforts to completely eradicate the spread failing, this no longer seemed like a feasible measure. Instead, many countries have turned to community mitigation strategies that help slow the spread of the disease in order to not overwhelm their public healthcare systems. These include but are not limited to wearing masks in public places, staying home when ill or displaying symptoms, or social distancing, where the public limits their nearby exposure to each other by avoiding gatherings or close contact.

Traditionally, general public's use of masks was *not* recommended because of a lack of data constituting mask usage among asymptomatic or healthy personnel (Aledort et al., 2007, CDC, 2019). Initially, many experts skeptically believed that the public would not be receptive to mask adherence and other social policies. Nevertheless, recent research suggests mask-wearing as an effective approach to containing COVID-19. For example, Dasgupta et al. (2021) pointed out that “The probability of a county becoming a rapid riser during the summer months was 43% lower among counties in states with statewide mask mandates at reopening… this association was more pronounced in non-metropolitan areas.” Despite being in the early stages of reopening the country, mask-abiding counties seemed to quell coronavirus outbreaks more effectively, with an immense impact factored into less densely populated locales. This may also be in part due to the general population being more mobile and less huddled indoors as temperatures rise in summer, lessening the potential for a mass outbreak of infections through direct contact. Moreover, among the other effective COVID policies, the mask mandates have fewer adverse effects on the economy and society if deployed earlier and implemented successfully (Krishnamachari et al., 2020). Zero-COVID functions involved the shutting down of non-essential public spaces and businesses through stay-at-home orders while limiting in-person interaction within the localized area with mass gathering closures, straining the workforce and community at large. Their effects took longer to noticeably reduce incidence rates and had clearly negative side effects ranging from economic output to mental wellness to academic performance. There must be other methods used in combating coronavirus without leaving incorporeal devastation across communities like a mask mandate.

Due to the different economic structures, city planning, and community mask-wearing behaviors across the world, the conclusions based on particular countries or states may not be applicable to one another. For instance, the effectiveness of the mask mandate in Australia may not be the same as in other countries, as these territories and provinces are not representative of US counties (An et al., 2021). Australia had contrasting collaborative methods and territorial divisions that resulted in differing resource allocations. Counties will have varying medical capacities, population health risks, or environments that may contribute to the incidence of disease. This research will substantially focus on Michigan as a separate territory and provide recommendations to public health officials, tailored specifically to Michigan.

With the ongoing COVID-19 pandemic, the public has been subject to differing recommendations from local, state, and federal officials and organizations. The WHO (World Health Organization) has recommended mask usage to prevent transmission and compensate whenever social distancing becomes impractical in public settings (as cited in Chaabna et al., 2021). This project will illuminate the efficacy of mask mandates within Michigan and what other recommendations health officials could make.

As an observational study where no randomized clinical trials are imposed on subjects, cause and effect are considered “speculative”, but generate associations that allow us to run future experiments on our hypotheses to be proven true. In the second section, we will investigate the claim that masks actively reduce the spread of COVID-19. In the third through fifth sections, we will measure the rates of disease incidence between counties that implemented mask-wearing requirements in public and private spaces compared to the times they were only recommended and discover whether any other variables affect the spread of COVID-19 within Michigan besides mask mandates.

Investigating the claim will allow us to understand the scientific rationale of the mask mandate. This would require reviewing any supporting literature. We will have to assess the efficacy of a mask mandate at the county level with repeated measurement and confounders adjusted. Finally, we give any appropriate recommendations for containing COVID outbreaks, specifically to Michigan counties. Moreover, we explore if any variable significantly affects Michigan and would alter COVID guidelines to tailor to the state's uniqueness.

2. Literature Review

In this time span, an enormous volume of scholarly literature has been published on disease transmission and how we can preventively counteract its spread.

Through use of a comprehensive meta-analysis over the breadth of material published in the last decade, research reviewers discovered that mask usage was associated with a general reduction in the contraction of respiratory viruses at an odds ratio of 0.35 (Liang et al., 2020). This applied to the populace not employed in a healthcare setting, demonstrating the generalizability of masks having a protective effect on all individuals, as masks halt the absorption of droplets or aerosols. For this study, it is assumed that all masks used by the public were surgical masks or N95 (US) or P2 (Australia) to cover their face. We look to see what face coverings factor into the protective effect.

The most commonly produced were surgical masks, face coverings that are obligatorily worn during medical procedures in order to not further contaminate the patient. They proved to be effective in having much lower viral loads detected compared to the subjects without masks (Leung et al., 2020). Masks lower the amount of viral shedding into the environment, thus reducing the spread and distance traveled of airborne breath particles. As a substitute, homemade or store-bought cloth masks were used when surgical masks were in short supply and stockpiled by federal authorities. However, they do not seem to provide the same general security as the aforementioned surgical masks. In a systematic search of research done on mask efficacy by Delvaeux et al. (2021), cloth masks generally did not offer the same defense as surgical and N95/P2 masks. There is a high degree of variability in how they were sewn, layered, and the materials used to produce them. Certain materials in these masks, such as cotton, are more permeable, allowing the penetration of airborne particles and/or droplets through the covering and increasing the chance of infection as they build up. Having more layers would increase the thickness, disallowing packs of particles to pass through the covering. In short, there is no easy way to judge their merits collectively with different standards. Therefore, our study will generalize to all masks, as Michigan requires any type of face covering. The effect of the mask mandate in Michigan might not equate to the efficiency of surgical masks as we focus on public policy.

A setback in these studies may be the issue of user compliance, where some of the experimental units may renege on their agreement as subjects randomly selected to their following experimental group, control or experimental, by not following instructions to wear a mask correctly (over the chin and nose) at all social instances. This affects the results by giving a false impression of the discordance of masks, as even temporarily lifting a mask in a social setting for an extended period of time can expose an individual to transmission. To mitigate this scenario, MacIntyre et al. (2009) adjusted the covariates to include the test participants who truthfully adhered to their mask usage to discover that in a pandemic setting where the perception of risk to harm or hazard is heightened, there would be a 60% reduction in infection's probability. This comes from a 26% decrease initially discovered where they were confident that the true value of the effect of mask usage globally within a day of incubation is between 0.09 and 0.77 in about 95% of all cases before the multiplicative 60% decrease with ethical mask users is applied. Masks tend to reduce the spread of coronavirus, and we would like to do this in a wide-ranging public setting instead of an experimental one.

3. Methodology

We obtain most of the datasets from the CDC website regarding the infection rates per county and other tracker sites regarding mask mandate enactment dates and who it applies for.

3.1 Data Description

From the CDC search engine, we chose to highlight three unique datasets that pertain to the necessary components to oversee the following: state transmission rates, rates exclusive to county populations, and level of vaccination across the United States. We merge the mask mandate information with the daily infection rates at the county level. Mainly, the data is being merged accurately and without losing information by identifying through FIPS (Federal Information Processing System), numbers used to identify geographical locations, such as counties in our case. Furthermore, FIPS does not attribute any similarly named counties that belong to Michigan to other US states.

Timewise, we focus on the data stream between summer 2020 and summer 2022, the beginning and end dates of enforcement, which is only a recommendation nowadays, where statewide mask mandates were most prevalent. The date is updated weekly as observations are routinely obtained in a timely manner. Our timeline is justified in starting from the beginning of the pandemic in March because data begins to be recorded by the states to trace the virus. Furthermore, key inflection points are the implementation of the mask mandate on July 10th, 2020, and the one-week early expiration of the state order on June 24th, 2021. Originally, it was set to expire at the start of July, but the governor allowed an early reprieve (Ainsworth, 2021). It lifts the remainder of restrictions and leaves recommendations up to officials and owners of public and private spaces, respectively, after the state order. The duration after the expiration of the mask mandate will be juxtaposed with the Biden administration's informal declaration of the pandemic's closure in September 2022. (Wolf 2022). This will determine whether there was any truth long after the mandates had concluded and the resumption of numerous daily tasks. Both periods, the activation and inactivity of the mask mandate, are roughly equal in length for one year, giving us balanced data for relatively unbiased results.

Vaccination data was initially added because it is a key component in the immunization process of the populace, especially in protecting the vulnerable elderly and youth's frail immune systems. Society would want to administer a safe, effective vaccine, but it will require time and patience to complete even an experimental version as soon as possible. We view vaccination as an important predictor in aiding the reduction in COVID spread. For measurement, there is a cumulative rate for complete vaccination through taking all necessary doses. For a maximum of 100%, all residents of an area are fully inoculated against the virus, thus reducing the chance of obtaining harsh symptoms that may hospitalize and reduce transmission. (CDC, 2022) As a precaution, the vaccine was not available to the public for all of 2020 and the first quarter of 2021, so the data will list the vaccination rate as 0%, hence delaying its effect.

In total, there are over 3000 counties that are recording over 130 weeks of data during this period. We will be dealing with continuous data as the uncountable range of the unknown distribution changes over time. As a result, discrete distributions are now exempt from our chosen models. If there was any missing or N/A data, the entire row would be automatically barred from any functions by R.

3.2 Linear Models

Normally, a linear model would be sufficient, but our response, what we measure, differs from counts, and the variance depends only on the mean. The ordinary least squares regression (OLS) does far less, since it assumes the relationship between our predictor(s) and responses is linear. We assessed the response, called the dependent variable, as the change in the infection rate per 100K over weekly intervals (per 100K implied whenever omitted). The focus on the change or difference in infection rate will allow use of the normal distribution. Otherwise, we have a zero-inflated model, a distribution with an abnormal amount of 0's (infection rate itself in this context) that skews heavily rightward. With higher values being a more drastic uptick or downtick in some cases, smaller spikes are more appealing as they demonstrate the spread is under the control of healthcare personnel. As the response variable involves proportions of repeated measurements, we suggest using the Generalized Linear Models (GLMs). It connects the conditional means of response to the linear predictors with the form:

$$g(E(y|x)) = \beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p$$

where

- y is the random response with the expected value conditioned on x
- β_0 represents the intercept or a constant value if the remaining predictors were 0 at the beginning, the superseding β's are regression coefficients
- *g()* represents the link function between predictors and the conditional mean of the response variable

The main assumption of the GLM family that is useful to us is the response belonging to the exponential family of the form: $f(y_i) = exp\left\{\frac{y_i\theta_i - b(\theta_i)}{a_i(\phi)} + c(y_i, \phi)\right\}$. Some examples are Gaussian (normal), Poisson (all positive values), gamma, etc. We do not know yet what distribution the model will take, so a preliminary analysis of the data needs to be completed in order to choose the best-suited distribution based on our knowledge of the data. Generalized linear models (GLMs) are a type of classical regression that determines the strength and nature of a statistical relationship between one or more independent variables and the dependent response, which allows response variables to have error distributions.

With various sources of variation within counties due to their structures and pandemic responses and our responses being repeated measurements, thus violating the assumption of independent observation, we should approach the problem with Generalized Linear Mixed-Effect Models (GLMM). GLMMs extend from the ordinary GLM model mentioned earlier, adding a "random effect" or random variable that follows $N(0, \sigma^2_a)$, where the deviation from the "fixed effect" is 0 and the variation occurs within its own nested group. For instance, these hierarchical groups in our study would be at the federal, state, or county level. Meanwhile, a fixed effect is the general level of variation without stratifying the data into subgroups; the same values as OLS. Fixed effects include accommodation of binary predictors like whether the mask mandate was implemented or not, while random effects allow accommodation across the differentiation in variation of geographic regions. In model estimation, the random effects are also known as random slopes that should sum to 0 in computation. Compared to classical statistical procedures, GLMMs provide a more flexible approach for nonnormal responses when temporal correlations are present. (Bolker et al., 2009). GLMMs combine the linear mixed models, which handle random effects, and the generalized linear models, which address the element of non-normality. In the research, GLMMs were able to handle the repeated weekly measurements that have various temporal correlation structures. We will implement the analysis with the R program (version 4.1.1) on Posit Cloud using the `lme4` package (Bates 2015).

4. Data Exploration

To find the overall change in infections, we should use the arithmetic mean as a measure of central tendency. It will average the change in infection rate for either the state or county levels over a certain duration. However, issues arise as places with lower populations tend to have lower infection rates because of the smaller pool of vectors for the virus to latch on to. There, we observe that a weighted mean is necessary, as population density fluctuates immensely between more compact city areas and sparsely inhabited rural areas. Without adjusting for population, the infection rate will face this undiagnosed confounder, skewing towards rural counties and proclaiming the mask mandate is not necessary in close-quarters urbanized areas.

Starting with a countywide exploration in Figure 1, the box and whisker plot is used to offer a visual of the distribution's spread through the five-point summary: minimum, 1st quartile (25th percentile), median, 3rd quartile (75th percentile), and maximum. In forming the axes, the side-by-side, vertical box plots consolidate the county's weighted mean for each month chronologically. We would like to see the change in COVID's infection rate by month within the pandemic's timeline. Figure 1 illustrates the IQR (Interquartile Range) expanding over the course of 2020 before deflating, where the median begins centering around the value 0 over the first half of 2021. Once again, this cycle repeats itself three more times, albeit to a lesser degree and in rapid succession. Many outliers exist beyond both whiskers. Nevertheless, the median converges to 0 as time goes on. As the president claimed that the pandemic was over, cases were transitioning into a decline in its third short cycle.

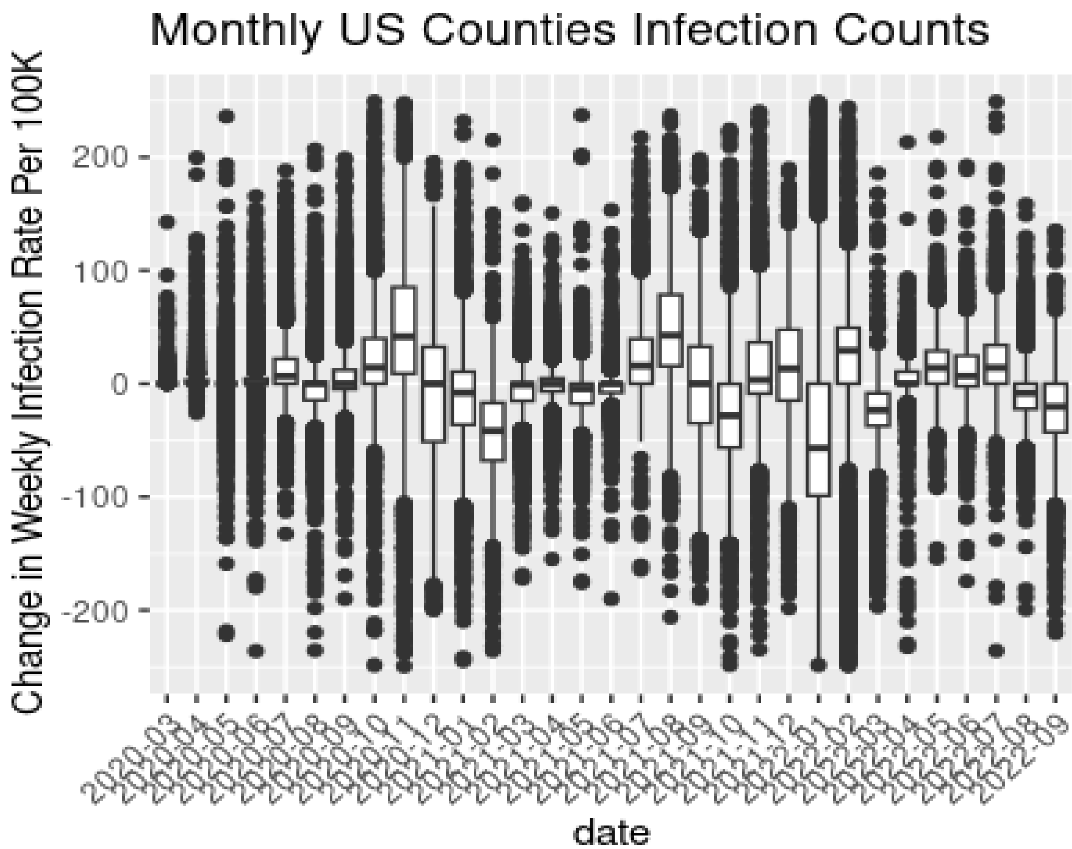

Figure 1: Box plot of US Counties Monthly Change in Infection Rate

From Figure 2 onward, any line plots will contain red dashed lines to signify the beginning and ending of the mask mandate. Figure 2 is a more localized approach from Figure 1, where the data is subsetted to only Michigan counties to investigate the trend of the change in the infection rate. Two small peaks of +100 cases pertain to the winter, while the other comes much later in mid-2021, not covered in the US counties in Figure 1. Again, the most rapid rise comes to fruition at the beginning of 2022. The weighted mean change in infections remains centered near the value of 0, with the worst rise being approximately +500 cases statewide. The takeaway is that there was little growth in infections during the mandate.

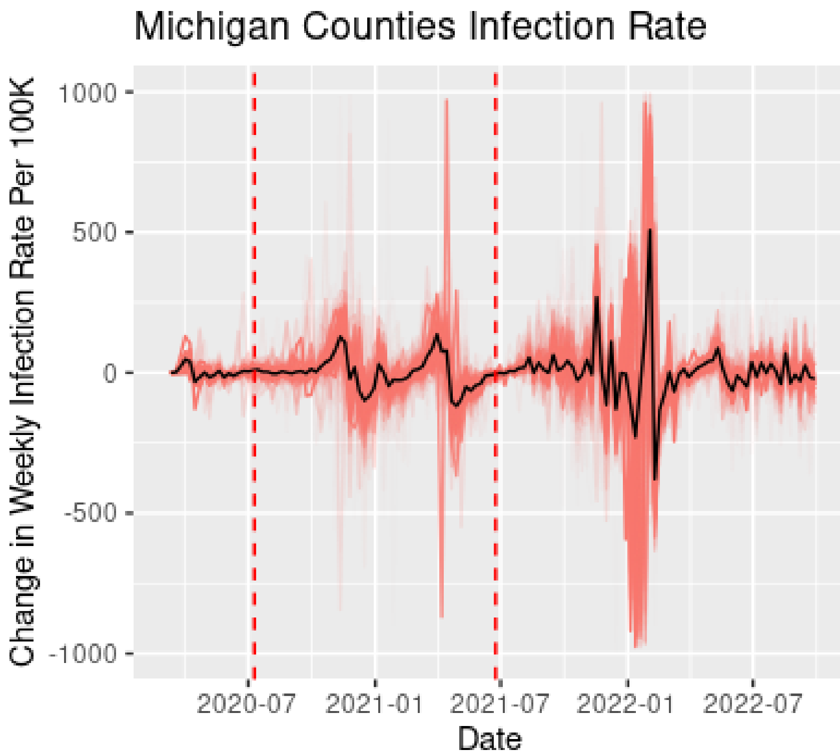

Figure 2: Trend of Michigan Counties's Change in Infection Rate

We note that in both Figures 1 and 2, the largest changes in the infection rate illustrate that winters would engender more illnesses due to closeness in proximity to others and lack of dispersion of the virus to the outdoors. Coincidentally, both correspond to the emergence of prominent COVID-19 variants. Variants are mutations from the virus's lineage that react to environmental concerns, affecting properties similar to spread, severity and development of new symptoms, and effectiveness of treatments. Specifically, Omicron became active in the United States in late 2021 with a more rapid spread than its less visible predecessor, Delta (Katella, 2023). The Delta variant had a more steady rise and fall of cases, where vaccines and immunity had their first adjustment period. We look to implement a binary predictor to highlight the timeline where they were the prevalent strain before vaccination changed and immunity heightened. Without them, our diagnostic models would show numerous outliers not accounted for during these rapid rises. They are listed as fixed effects due to their indiscriminate nature in spreading to all 83 counties of Michigan, overtaking the original lineage.

Next, we present how starkly different the change in infection rate was between some of the most and least COVID-stricken counties in Michigan and what could be attributed to this divide. In this graph, the worst counties in regard to COVID infection rates are marked in red, while the best counties are signified in blue. As shown below in Figure 3, the difference between the bottom ten and top ten worst performing countries is estimated to be a weighted mean change in 100 new infections. Beware, as this was not as substantial a difference as anticipated. There is quite a lot of overlap between both trends other than the peaks and troughs, so the difference between counties may be marginal, suggesting the random effects are not statistically significant. If further analysis proves this inference incorrect, then this provides evidence to the idea that counties are heterogeneous based on infection rate, so a mixed model would be appropriate to accommodate the issue at hand.

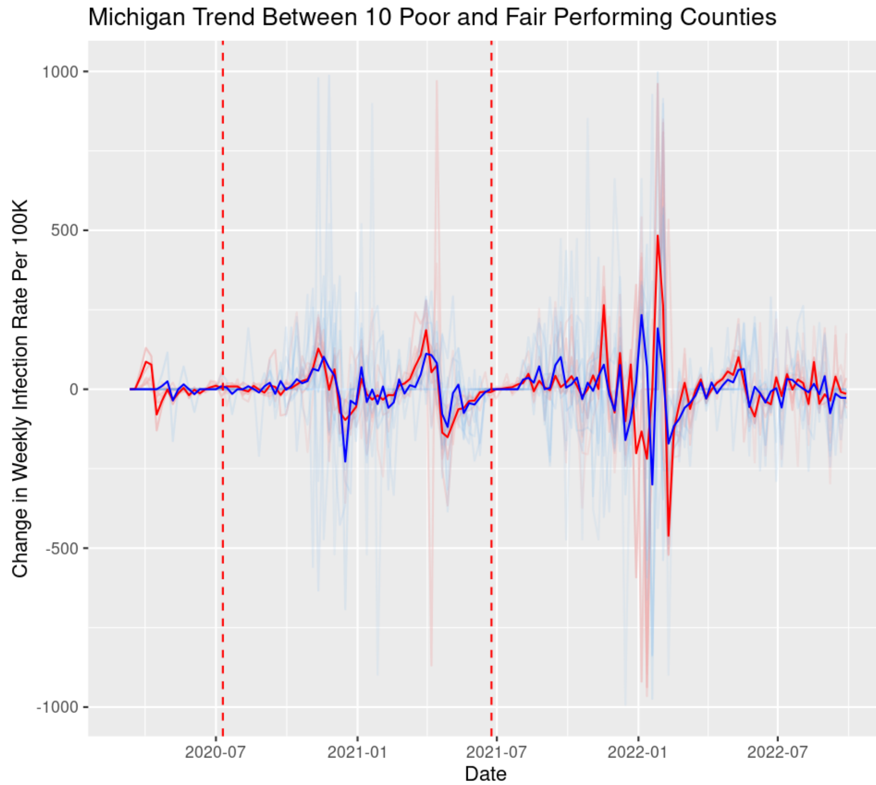

Figure 3: Top/Bottom 10 Trend Comparison of Mean Change in Infection Rate

To clearly visualize that there was a distinct change in infection rate in Michigan, we plotted maps onto three grids, featuring the weighted mean changes in infection rate at defined durations in Michigan counties. Figure 4 demonstrates that there is a change in the mean infection rates during these periods. The mean change in infection rates dipped during the mandatory period and then swelled once it was revoked. However, it would be illogical to say that factors other than the mask mandate are not contributing to the decline in case numbers. Therefore, we should study a model that takes other variables into consideration and adjusts for their effects.

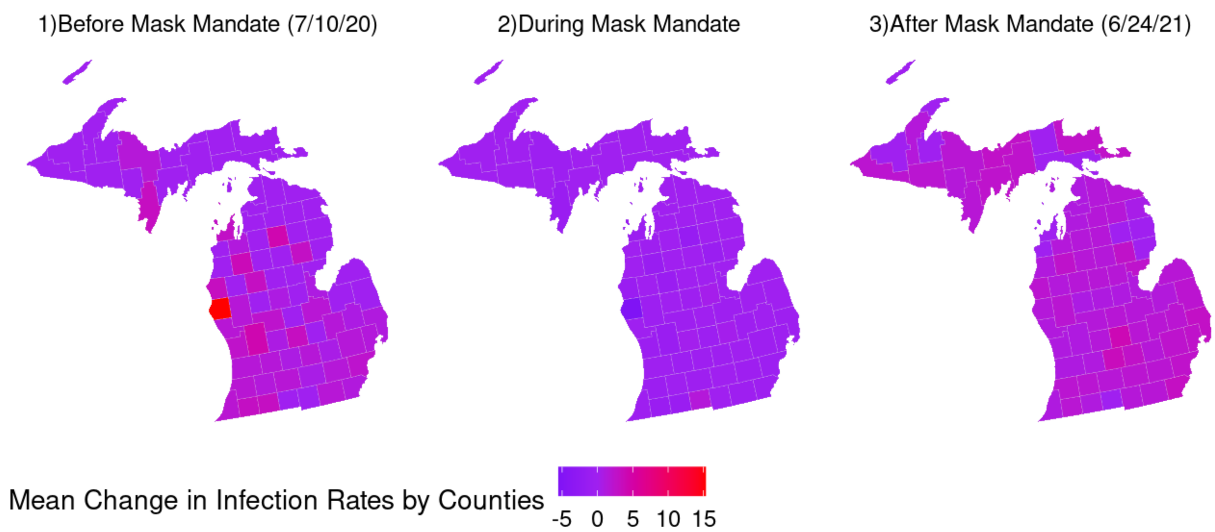

Figure 4: Michigan County Map Plots of the Mean Change in Infection Rates at Various Time Points

Recognizing that differences in settlement sizes may contribute to greater spread, population density requires consideration as well. Notably, there was no dataset distinguishing urban from rural that could be properly carried through, so it was manually coded for Michigan. To distinguish population density within Michigan, the US Census Bureau defines an urbanized county as a territory that holds at least 50,000 citizens collectively or adjacent territory that aids in the daily transportation of these surrounding citizens to and fro (State of Michigan n.d.). We should not overlook this potential predictor, even when already factoring county populations.

In creating our model, we must have an idea of what pool of variables will be impactful to the regression, as we do not want to overfit the model, resulting in a failure to forecast future values or make robust predictions. Previously, we named certain variables that are believed to predict the response. Now, we should configure a scatter plot matrix where the two-dimensional space contains all pairwise combinations in (x, y) coordinates. In Figure 5, we compute the relationship between any of the few numeric variables within our datasets. For starters, the first and booster doses are strongly correlated with the percent weekly vaccination. To avoid multicollinearity issues, we assign the percent of weekly vaccinations as the only variable regarding immunization to prevent erratic changes between the predictors themselves. The remaining variables passed the first screening, so more objective methods will be trialed.

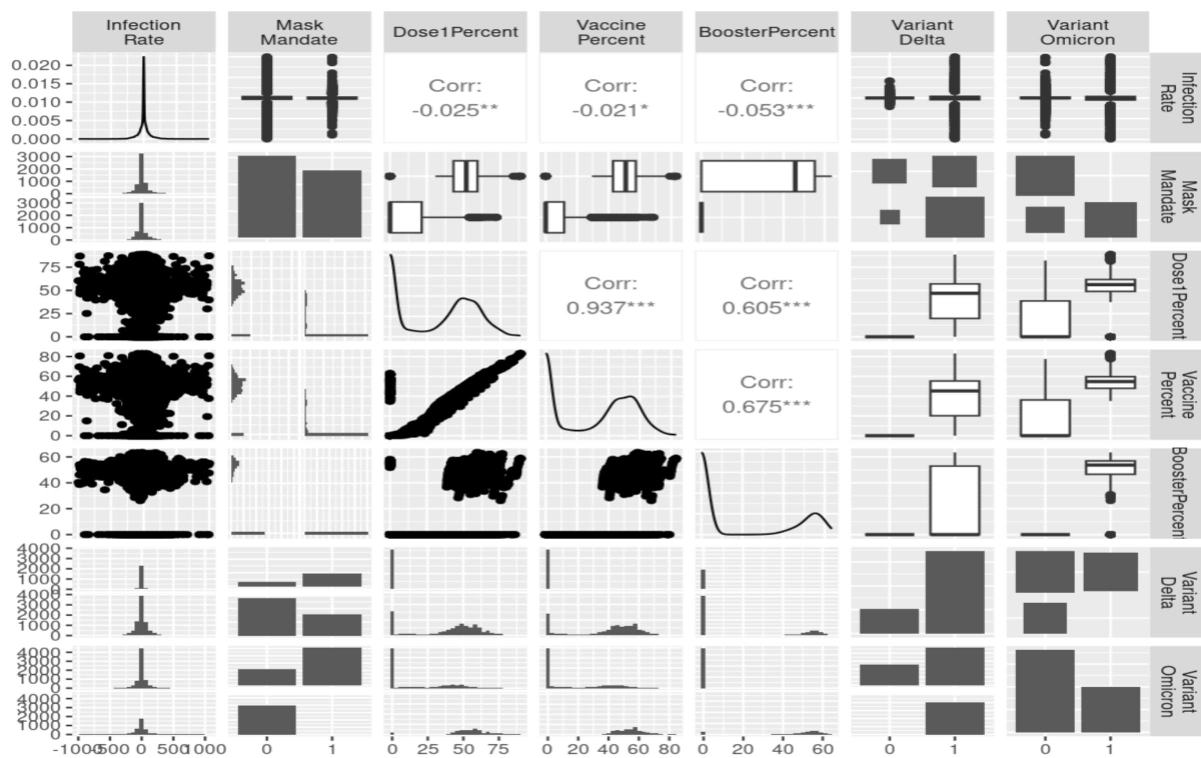

Figure 5: Relationships between All Explanatory Variables

5. Model Specification

Essentially, we seek a simple model with the least number of terms to accurately predict the weekly change in the COVID infection rate. Complex models with too many terms are ripe for overfitting and cannot extrapolate properly. Viewing a histogram of our response variable, it displays the probability distribution density (pdf) of the response variable, which belongs to the influx of positive and negative values. What we have is a roughly symmetric curve centered around 0, with large lagging tails at both ends from outliers. Following an identity link (normal distribution as default), a LMER model is fitted instead of the exponential family of GLM.

To test available variables for significance, we create nested LMER models to highlight chosen terms. When comparing a duo of models, a likelihood ratio test (LRT) is usually used, but the Kenward-Roger test is particularly useful in analyzing the LMER models to see whether a term in the function is significant enough to be kept. Our null hypothesis (H_0) is that the term is not necessary, and the alternative hypothesis (H_a) is that the term should be kept. It offers a p-value, the probability that the observation is seen as extreme, where being less than our set alpha level, such as default 0.05, would allow us to reject the null hypothesis with 95 % confidence that it was a correct decision without invoking type Ⅰ error (false positive). With the acceptance of the alternative, it would signal that one model is better than the other. There is a manner in which we can measure all potential models simultaneously.

We choose to use backwards elimination as a part of the stepwise approach. This method provides all the variables at the start and drops the lowest-performing terms repeatedly through a rigorous alpha level until the lowest AIC (Akaike Information Criterion) is found. Prior to initializing the test, we are forced to eliminate a random effect, as the program could not partake in all three without an eigenvector being unavailable.

Effects	**Degrees of Freedom**	**Sum of Squares**	**Residual Sum of Squares**	**AIC**	**F value**	**P-value**
Active Mask Mandate	**1**	**466257**	**192326515**	**108537**	**27.0164**	**2.053e-07**
Percent Weekly Vaccinations	**1**	**137400**	**191997659**	**108518**	**7.9614**	**0.004787**
Variant Delta	**1**	**66707**	**191926966**	**108514**	**3.8652**	**0.049321**
Variant Omicron	**1**	**274975**	**192135234**	**108526**	**15.9330**	**6.605e-05**

Table 1: Backwards Elimination Final Results

Table 1 shows the remaining terms not discarded from the repeated Satterwaite t-tests. They are all fixed effects with degrees of freedom showing only one level by definition. Sum of squares (SS) measures how well the term fits among the dispersion, and residual sum of squares (RSS) measures the error, ε, between the data and function of the regression model. High test statistics and low p-values exemplify significance predicting the response of the model. According to the tests, no random effects were considered remotely significant, as they were some of the first terms to be eliminated. The weekly culmination of vaccines is surprisingly less significant than either of the remaining three variables, although it passed many Kenward-Roger tests that the mandate failed. The random effect of the mask mandate was eliminated, but this was only a suggestion. Since we are measuring the random effect, it will be included in the final model along with a random intercept to exhibit variation exists at the county level.

6. Outcomes

We learned that the three fixed effects are crucial to fitting the model. Despite dropping the random effects in stepwise regression, we still believe that they are important to understand if there are differences between the counties. Here is our finalized model:

Mean of the Change in Infection Rates Per 100K = (1+Active Mask Mandate | County) + Active Mask Mandate + Percent Weekly Vaccinations + Variant Omicron + Variant Delta

*Note: **(1+Active Mask Mandate | County)** is a random effect applied to both terms*

Fixed Effects	Estimate	Standard Error	T Value
Intercept	**19.0646**	**4.1566**	**4.587**
Active Mask Mandate	**-21.8071**	**4.1955**	**-5.198**
Percent Weekly Vaccinations	**-0.2946**	**0.1044**	**-2.822**
Variant Omicron	**-15.9174**	**3.9877**	**-3.992**
Variant Delta	**7.8819**	**4.0091**	**1.966**

Table 2: Fixed Effects of Model Equation

Based on Table 2, the mask mandate's fixed effect suppresses coronavirus cases by an estimated 22 per 100,000 people. Every 1% increase in vaccination prevents coronavirus cases by an estimated .3 per 100K. Vaccination efforts could gain a similar negative coefficient as the mask mandate, at approximately 70% vaccination rate. The variants effects may be misleading as Omicron rose aggressively in case values while Delta steadily increased over a quarter of a year before the first major changes in protocol were initiated.

Groups	**Name**	**Variance**	**Standard Deviation**	**Correlation**
County	**Intercept**	**3.937e-09**	**6.274e-05**	
	Active Mask Mandate	**7.035e-09**	**8.388e-05**	**-1.00**
Residual		**1.726e+04**	**1.314e+02**	

Table 3: Random Effects of Model Equation

In line with Table 3, the fixed effects standard errors from Table 2 change when in tandem with the random effects. Unfortunately, this change is negligible, as the random effects' variances are too low to have any effect on our model. In particular, manually adding the random effects back still shows that the county differences are not statistically significant in any form. Regarding AIC, 140,065 is our information criterion, an extremely high value due to the amount of data points, but nonetheless pared down from around 150,000 using the original response, infection rate per 100K.

Checking the model diagnostics in Figure 6, we first look at a Q-Q (quantile to quantile) plot, a scatter plot between theoretical and actual percentiles, to check the assumption of normality. It looks to be a heavy-tailed distribution featuring two small "bends", suggesting nonlinearity. Most likely, higher-order terms like a cubic function or interaction between predictions may be added to attempt a straighter line. Some outliers lie outside the model's fit, as underscored by the red points. Yet, Cook's distance, an estimate of the influence of a point, provides us a visual of where much of the data is below the cut-off region, influencing the fitted data only sporadically. The variance from the scale-location plot, commonly called the residual plot, shows a roughly homogeneous spread where little discernible shape crops up other than at the fitted value of 5, where low vaccination and Delta occurs. This is a well-fit model.

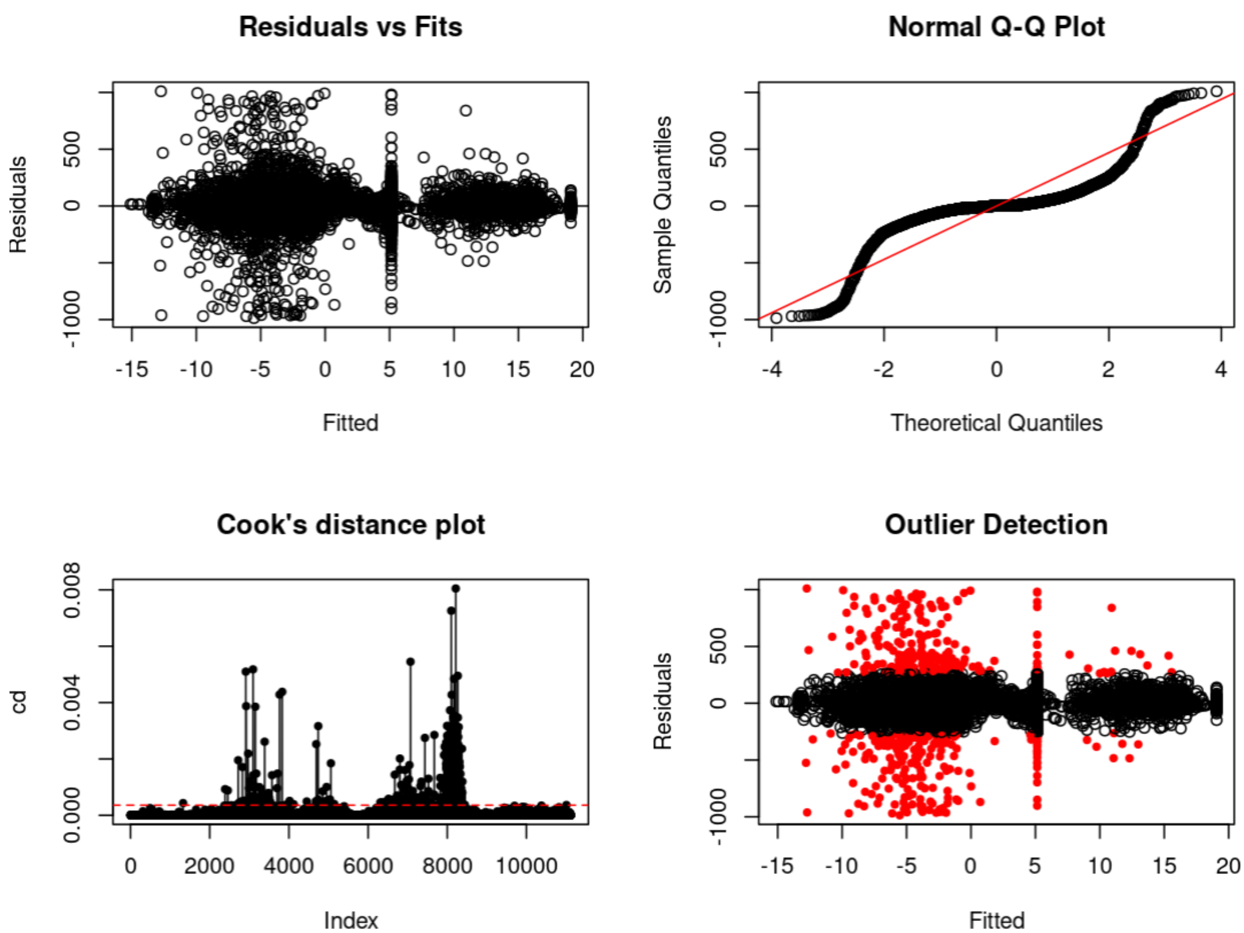

Figure 6: Model Diagnostic Plots (2x2)

7. Conclusion

As mentioned in (Adjodah et al., 2021), masking has been proven to reduce the spread of coronavirus among the general worldwide population and is expected to be efficacious in Michigan as well. More importantly, masking limits the spread of influenza diseases that have the same mechanisms as coronavirus. The fixed effect of the mask mandate is effective at the state level of Michigan while the random effects are trivial. Between counties, there was little to no efficacy concerning the spread. Most importantly, the process of gathering these results will be replicable to draw from other surrounding territories in the future, barring our limitations.

Many countries in the western hemisphere have not been ardently supportive of mask mandates. Specifically, they value individualism and making their own choices as freedom and civil liberty, whereas the mandate is seen as an encroachment on their lives. To achieve better mask-wearing behavior, different cultures can foster different values. For example, Asian societal norms when one is sick expects the strict wearing of a mask. To evaluate, Liang (2020) discovered an even greater protective effect for masks generated at a 69% decrease in infections in Asian countries, whereas western countries had only a 55% decrease. With Asia's countries holding more attitudes that value the collective good of society as a whole, there is little argument that their methods are much stronger. The United States and other applicable western countries should look to gradually adopt these norms through the proliferation of PSAs (Public Service Announcements) and alterations to public guidelines to better match their healthier counterparts.

An improvement would be to model the data as a time series. A time series involves the measurement of a dependent variable while a time index is chronologically followed. We are able to use the property of stationarity, satisfying the conditions regarding independence of the mean function and time, to discover the trend through a residual process. Time series analysis allows us to not only measure the trend in cases but also whether there is a factor of seasonality, more so, whether caseloads change with the seasons like winter. Furthermore, the volatility in the heteroscedastic variance of the continuous data alongside tail heaviness may offer a solution through the use of a GARCH (Generalized Auto Regressive Conditional Heteroskedasticity) model. This allows the error term variance to be correlated with the previous error size. Refer to Figures 2 and 3 to view fluctuations in variance that GARCH may alleviate.

Another limitation we encountered was the compilation of the mask mandate dataset becoming obsolete at the county level. While statewide mask mandates were more readily easy to gather and record, researchers could not find the manpower or time to track where some counties' mask mandates began or finished. Other problems persisted from the police departments; non-enforcement of the law based on the sheriff’s decision would dissuade immediate implementation, while residents would not follow the emergency law without significant penalty, leading to an early repeal of the orders, both statewide and countywide. Ultimately, we went on a different trajectory with more accurate data and a slightly modified approach to the issue of mask mandate policy.

Lastly, we originally assumed the citizens of each county do not travel often or do not bring the infection to other locales. In reality, this proves to be demonstrably false. Basically, the infected population traveled around post-lockdown, typically into adjacent counties if they lived near the border of a county. For instance, Michigan is divided into subregions, such as the Mackinac Bridge being a man-made barrier between the Upper and Lower Peninsulas. We could code for different regions of Michigan by combining counties to investigate their relationship regarding infection rates. In addition, the Midwest, a region including Ohio, Illinois, Wisconsin, Minnesota, Indiana, could have compared its mask mandates with the rest of the country due to their shared characteristics.

Bibliography

Adjodah, Dinakar, K., Chinazzi, M., Fraiberger, S. P., Pentland, et al. (2021). Association between COVID-19 outcomes and mask mandates, adherence, and attitudes. *PloS One*, *16*(6), e0252315–e0252315. https://doi.org/10.1371/journal.pone.0252315

Ainsworth, A. (2021, May 21). When does Michigan's mask mandate end? *FOX 2 Detroit.* Retrieved from https://www.fox2detroit.com/news/when-does-michigans-mask-mandate-end.

Aledort, J.E., Lurie, N., Wasserman, J. et al. (2007). Non-pharmaceutical public health interventions for pandemic influenza: An evaluation of the evidence base. *BMC Public Health* **7,** 208. https://doi.org/10.1186/1471-2458-7-208.

An, B.Y., Porcher, S., Tang, S., & Kim, E. E. (2021). Policy design for COVID‐19: worldwide evidence on the efficacies of early mask mandates and other policy interventions. *Public Administration Review, 81*(6), 1157–1182. https://doi.org/10.1111/puar.13426.

Bates, D., Maechler, M., Bolker, B., Walker S., (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1-48. doi:10.18637/jss.v067.i01.

Bolker, B. M., Brooks, M. E., Connie, J. C., *et al.* (2009). Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in Ecology & Evolution*, *24*(3), 127-135, https://doi.org/10.1016/j.tree.2008.10.008.

Centers for Disease Control and Prevention (CDC). (2022). COVID-19 vaccination. Retrieved 2022 from https://www.cdc.gov/coronavirus/2019-ncov/vaccines/different-vaccines/overview-COVID-19-vaccines.html.

Centers for Disease Control and Prevention (CDC). (2019). Interim guidance for the use of masks to control influenza transmission. Retrieved 2022 from http://www.cdc.gov/flu/professionals/infectioncontrol/maskguidance.htm.

Chaabna, Doraiswamy, S., Mamtani, R., & Cheema, S. (2021). Facemask use in community settings to prevent respiratory infection transmission: A rapid review and meta-analysis. *International Journal of Infectious Diseases*, *104*, 198–206. https://doi.org/10.1016/j.ijid.2020.09.1434.

Chung, Marlow, S., Tobias, N., Alogna, A., Alogna, I., You, S.-L., Khunti, K., McKee, M., Michie, S., & Pillay, D. (2021). Lessons from countries implementing find, test, trace, isolation and support policies in the rapid response of the COVID-19 pandemic: a systematic review. *BMJ Open*, *11*(7), e047832–e047832. https://doi.org/10.1136/bmjopen-2020-047832.

Dasgupta, S., Kassem, A. M., Sunshine, G., *et al.* (2021). Differences in rapid increases in county-level COVID-19 incidence by implementation of statewide closures and mask mandates — United States, June 1–September 30, 2020. *Annals of Epidemiology*, *57*, 46–53. https://doi.org/10.1016/j.annepidem.2021.02.006.

Delvaux, Aertgeerts, B., & Verbakel, J. Y. J. (2021). BET 1: Do homemade or cloth face masks work as a preventive measure for respiratory virus transmission? *Emergency Medicine Journal : EMJ*, *38*(5), 401–403. https://doi.org/10.1136/emermed-2020-209761.2

Faraway, J. J. (2016). Extending the linear model with r: Generalized linear, mixed effects and nonparametric regression models. *Chapman & Hall/CRC.*

Fox, J., & Weisberg, S. (2019). An R companion to applied regression. *SAGE.*

Katella, K. (2023, February 3). Omicron, Delta, Alpha, and more: What to know about the coronavirus variants. *Yale Medicine.* Retrieved from https://www.yalemedicine.org/news/covid-19-variants-of-concern-omicron.

Krishnamachari, Morris, A., Zastrow, D., Dsida, A., Harper, B., & Santella, A. J. (2021). The role of mask mandates, stay at home orders and school closure in curbing the COVID-19 pandemic prior to vaccination. *American Journal of Infection Control*, *49*(8), 1036–1042. https://doi.org/10.1016/j.ajic.2021.02.002.

MacIntyre, Cauchemez, S., Dwyer, D. E., Seale, H., Cheung, P., Browne, G., Fasher, M., Wood, J., Gao, Z., Booy, R., & Ferguson, N. (2009). Face mask use and control of respiratory virus transmission in households. *Emerging Infectious Diseases*, *15*(2), 233–241. https://doi.org/10.3201/eid1502.081167.

Midway, S. (2022). Data analysis in R. Retrieved February 5th, 2023 from https://bookdown.org/steve_midway/DAR/.

State of Michigan. (n.d.). Appendix b: Rural and urban county classification. Retrieved from https://www.michigan.gov/-/media/Project/Websites/mdhhs/Folder2/Folder48/Folder1/Folder148/B-Urban_Rural.pdf?rev=f376559abd34451db54fb72ea120019e.

Wolf, Z. B. (2022, September 20). Biden declares the pandemic over. People are acting like it too. CNN. Retrieved from https://www.cnn.com/2022/09/19/politics/biden-covid-pandemic-over-what-matters/index.html.

Appendix

R code file for reproducibility: PiwowarskiD/PiwowarskiHCThesis: Michigan/US COVID Data (github.com)

Warning: Must have R installed on personal device by https://www.r-project.org/

Datasets for R Folder:

COVID-19 Vaccinations in the United States,County | Data | Centers for Disease Control and Prevention (cdc.gov)

Weekly COVID-19 County Level of Community Transmission Historical Changes | Data | Centers for Disease Control and Prevention (cdc.gov)

United States COVID-19 Community Levels by County | Data | Centers for Disease Control and Prevention (cdc.gov)

https://www.austinlwright.com/covid-research.

Shiny: https://piwowarskidavid.shinyapps.io/USCOVIDInfectionRates/