**COVID-19 IN MARYLAND**

DATA SET: <https://covidtracking.com/data/download>

DATA DEFINITIONS: <https://covidtracking.com/about-data/data-definitions>

**12/09/2020**

**HIT 750 DATA ANALYTICS**

UMBC Fall 2020

Group 5

**Hala Algrain, Eric Kelly, Zenas Njokom, Rohini Salla**

**Abstract**

Since the first cases of COVID-19 were reported in late 2019, countries have attempted to limit the spread of the disease using different methods. The purpose of our research is to identify the effectiveness of some of these different methods. To do this, we examined the mandates and restrictions issued by the state of Maryland. The data that we used includes: positive increase, death increase, hospitalization increase, and recovered. We then compared this data to the timeline of restrictions and mandates imposed by the Maryland government. Our results showed that there is a significant difference in the rate of COVID cases between the successive months of March through July (p<0.01), but no significant difference between the successive months of July through October. Our data suggests that mandates and restrictions, alongside levels of enforcement and community behavior patterns, have an impact on COVID-19 rates.

**Introduction**

Since the first reported cases of COVID-19 were reported in late 2019, doctors and scientists have been brainstorming ways to limit the spread of the disease. On 23 January 2020, the central government of China imposed a lockdown in Wuhan and other cities in Hubei in an effort to quarantine the center of an outbreak of coronavirus disease 2019 (COVID-19); this action is commonly referred to as the Wuhan lockdown. Once the disease spread to other countries, worldwide social distancing and national quarantines ensued. As a result of this, we experienced one of the quickest recessions in our nation’s history. Unemployment skyrocketed, businesses closed, unessential travel ceased. However, how much of an impact have these precautionary measures in limiting the spread of COVID?

Just as countries have adopted different strategies to combat COVID so has the United States. Some states have had loose restrictions while others have taken aggressive steps in an attempt to limit the transmission of the disease. These different strategies are necessary because of the different population density of the states; what may work in Wyoming may not be the best approach in New York City. The strategy that Maryland has taken has loosely coincided with recommendations from the CDC and the current administration. Since the first COVID case in Maryland, we can look at how these restrictions have impacted the spread of the disease.

The best source that we have found with the most comprehensive dataset of COVID tracking data is from The COVID Tracking Project. Our group will analyse the most pertinent statistical COVID related information in the dataset. We will then compare the timeline of restrictions and mandates to the dataset. The goal of our data analysis is to determine the efficacy of government mandates/restrictions to combat the spread of COVID-19 in Maryland.

The goal of our data analysis is to determine the efficacy of government mandated sanctions to combat the spread of COVID-19 in Maryland. The data set that we will use is from The COVID Tracking Project. This provides the number of COVID cases and number of COVID deaths in Maryland, along with other relevant statistics. We will compare the prevalence of the disease with the timeline of government mandated restrictions and advisories. We will then be able to see how effective these restrictions are at limiting the spread of the disease.

We hypothesize that the restrictions are having an effect on the spread of COVID. However, we do not know how big of an effect it is having. Since we do not know, we are very interested to complete our research to find which restrictions/mandates are effective, and which are not. Since the beginning of COVID, we have already learned that some precautions are effective. In March, Dr. Fuaci said that masks were not recommended, later, the CDC stated that masks are strongly recommended. Then many governors, including Maryland governor, Larry Hogan, made masks mandatory. Our goal is to explore whether the different lockdown measures have had any impact on the number of positive cases, hospitalizations, and/or deaths.

***Null Hypothesis***; the lockdown measures mandated in Maryland had no significant impact on the spread or number of COVID-19 cases witnessed.

***Alternative Hypothesis***; the lockdown measures mandated in Maryland had a significant impact on the spread or number of COVID-19 cases witnessed.

Hypothesis testing: from our analysis, we shall then determine if the rate of positive cases increase, actually increased or decreased due to the mandated lockdown measures. Depending on our findings, we shall then decide to retain the Null hypothesis or Alternative hypothesis and which to reject.

**Related Work**

COVID-19 is a type of virus that is similar to the common cold. It was first discovered in humans in the 1960s. Since then, it has been studied in several labs around the world to better understand the virus, and further vaccine research.

Face coverings- Early in the COVID-19 pandemic, the WHO, the CDC and NIH’s Dr. Anthony Fauci discouraged wearing masks as not useful for non-health care workers. Now they recommend wearing cloth face coverings in public settings where other social distancing measures are hard to do.

According to The Association of American Physicians and Surgeons, “Surgical masks are loose-fitting devices that were designed to be worn by medical personnel to protect accidental contamination of patient wounds, and to protect the wearer against splashes or sprays of bodily fluids. They aren’t effective at blocking particles smaller than 100 μm”. This is a problem because COVID is approximately 0.12 microns. They were designed to protect against droplets, not aerosols.

Are people wearing masks correctly? In a study in Singapore, data was collected from 714 men and women. Of all ages, only 90 participants (12.6%) passed the visual mask fit test. About 75% performed strap placement incorrectly, 61% left a “visible gap between the mask and skin,” and about 60% didn’t tighten the nose-clip. Masks, which are already not very effective against spreading COVID, are nearly useless if not worn properly.

Social distancing- How effective is social distancing and where did the “6 foot” rule come from? Large respiratory droplets (>5 μm) remain in the air for only a short time and travel only short distances, generally <1 meter. They fall to the ground quickly. This idea guides the CDC’s advice to maintain at least a 6-foot distance. Larger particles land on nearby surfaces quickly, anything over 500-micron will be airborne less than 1 second.

How far does COVID-19 travel in air? According to Renown Health Products, “A 100-micron particle will fall for about 6.7 seconds in still air can travel about 6ft with no assistance from additional airflow. Particles less than 100-micron have the potential to be entrained in the airflow pathway for some time and particles <50-micron may not settle, dependent on particle size and indoor airflow. Particles smaller than 7-micron are easily entrained in the airflow pathway, rather than settling to the ground”. They can stay suspend in air and replication-competent for extended periods, therefore, social distancing will not mitigate risk from particles this size. This data suggests that close contact with a symptomatic person, even while wearing a surgical mask is not an effective barrier. However, simply passing a person with symptomatic COVID will not put you at a high risk of transmission.

Many different organizations have tracked COVID in countries around the world. They have put the data into line graphed so we can visualize trends. One organization that has done this is WorldMeter. From this, we are able to visually compare the countries’ COVID rates and tell if their mandates and restrictions are working. For the most part, they all follow a similar trend- peak in early spring, drop during the summer, and a second wave in the fall/winter months.

**Methodology**

The data used in this study came from the COVID Tracking Project, which is a volunteer organization launched from *The Atlantic* and dedicated to collecting and publishing data tracking COVID-19 outbreak throughout the United States. Data on COVID-19 testing and patient outcomes from all 50 states, 5 territories, and the District of Columbia were collected on a daily basis. Most of the data compiled were taken directly from the websites of local or state/territory public health authorities. In a case where data were missing from these websites, the missing information was supplemented with available numbers from official press conferences with governors or public health authorities. The website contains data from March 14 to date. For the purpose of this study, we limited our scope to Maryland, by examining reported data from March 14 through October 20. The dataset contained the following columns:

date, state, positive, positiveIncrease, positiveCasesViral, negative, negativeTestsViral, pending, positiveTestsViral, totalTestsPeopleViral, totalTestsViral, totalTestEncountersViral, negativeTestsPeopleAntibody, negativeTestsAntibody, positiveTestsPeopleAntibody, positiveTestsAntibody, positiveTestsPeopleAntigen, positiveTestsAntigen, totalTestsPeopleAntigen, totalTestsAntigen, hospitalizedCumulative, inIcuCumulative, onVentilatorCumulative, hospitalizedIncrease, death, deathConfirmed, deathProbble, deathIncrease, recovered, dataQualityGrade.

See appendix B for more description of each variable. Furthermore, for the purpose of this study, we considered only certain variables that were found important after performing some basic data wrangling (cleansing) as seen in the results.

We used R-Studio (Version 1.3.1073) as the integrated development environment for statistical computation, analysis, and visualization of our data.

The reference periods we tried for estimating the effects of lockdown mandastes was available data prior to March 16, 2020 when we had the first mandate that called for public schools to close. Also, the time frame between each mandate served as some reference to preceding the preceding executive mandates that were signed since many mandates were ordered at different times. We examined how effects change over the following dates :

March 16, 2020: Schools closed, executive order closes public places

March 23, 2020: Hogan ordered nonessential businesses to close

March 30, 2020: Governor issued stay-home order

April 15, 2020: Hogan signed face-mask order

May 13, 2020: Hogan announced Stage One of reopening (effective 15 May)

July 29, 2020: Hogan expanded mask order, issues out-of-state travel advisory

Aug. 27, 2020: All schools authorized to reopen, Hogan says

Sept. 1, 2020: Maryland entered Stage Three of recovery plan.

Sept 21: The governor announced restaurants could expand indoor capacity from 50% to 75% beginning at 5 p.m.

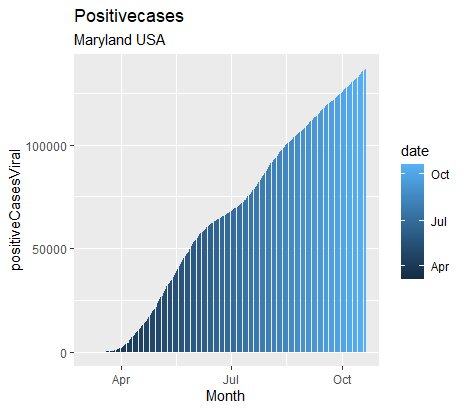
**Statistical Analysis**

We referenced the stratified sampling model, to examine whether statewide mandates actually had an impact on the spread of COVID-19 in the state of Maryland. This method allowed us to divide the total population-data into smaller groups or strata (months / periods related to mandate dates) to better understand existing relationships and trends between different sampling groups (months, weeks or days) and compare the pre-post mandate changes of COVID-19 spread in Maryland over time. By using a stratified method, we were able to examine the timeline of restrictions and mandates, and if the imposed mandates had an impact on the spread of COVID-19 in Maryland.

After we categorized into months, we plotted regression lines to see if months had different trends. We then analyzed the data using one-way ANOVA and subsequent to finding statistical significance we used a Tukey Test to find out which of the months were statistically significant from one another.

**Results**

**a)** We looked into the number of positive cases viral from March through October. Positive cases viral is defined as the total number of unique people with a positive PCR or other approved nucleic acid amplification test (NAAT), as reported by the state or territory. This is equivalent to a confirmed case as per the Council of State and Territorial Epidemiologists (CSTE) case definitions.



*The above graph shows a bar chart of Positive cases viral over a period of March to mid October.*

As shown on the positive cases viral graph, Maryland appears to have a worse trajectory, showing an average increasing trend with multiple spikes in COVID-19 cases.

However, Maryland has started restricting businesses, closing schools, and promoting social distancing a bit earlier than others relatively. Lockdowns don’t cover the entire state. There’s a patchwork of different restrictions by city and county. So, the numbers are trending upwards and the graph shows an increasing trend.

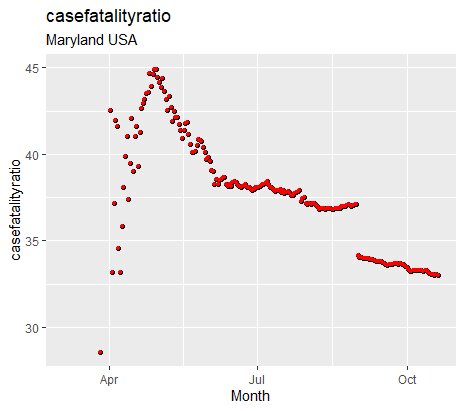
**b)**. Another interesting information that was worth investigating was the case fatality ratio**.** *The case fatality ratio is the proportion of individuals diagnosed with a disease and ended up dying from that disease.* Thus, a measure of severity among detected cases*.*

According to WHO, the Case Fatality ratio can be calculated by the below formula for an ongoing pandemic.

**Case fatality ratio (CFR) =**

**(Number of deaths from disease/(Number of deaths from disease + Number of recovered from disease))\*100**

Although the case fatality is not a biological constant, nevertheless it is useful to represent the magnitude of the disease at a given time, in a given population in a specific context.



*The above scatter plot shows the case fatality ratio for COVID-19 in Maryland over time, 20 October 2020.*

From the above graph, interpretation is that the CFR was much higher in the earliest stages of the outbreak. But in the weeks that followed, the CFR declined, reaching as low for patients who first showed symptoms.

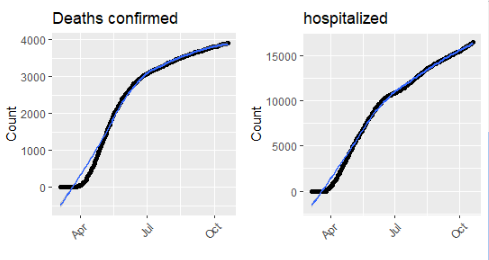
This is because, according to the WHO, "the standard of care has evolved over the course of the outbreak."

CFR can decrease or increase over time as responses change and differ according to the infected population's location and its characteristics such as age or gender. Older people, for example, are projected to see more COVID-19 CFR than younger populations.

Strict lockdown strategies together with a wide diagnostic PCR testing of the population were correlated with a relevant decline of the case fatality ratio.

**c)**. Deaths confirmed and hospitalized totals were graphed separately. These graphs tracked the total deaths among confirmed COVID-19 cases and hospitalized cases in Maryland over time. The data points on the deaths confirmed graph represent the cumulative total number of deaths reported to public health by the date along the bottom.This graph presents data by the date a death was reported as being associated with COVID-19.

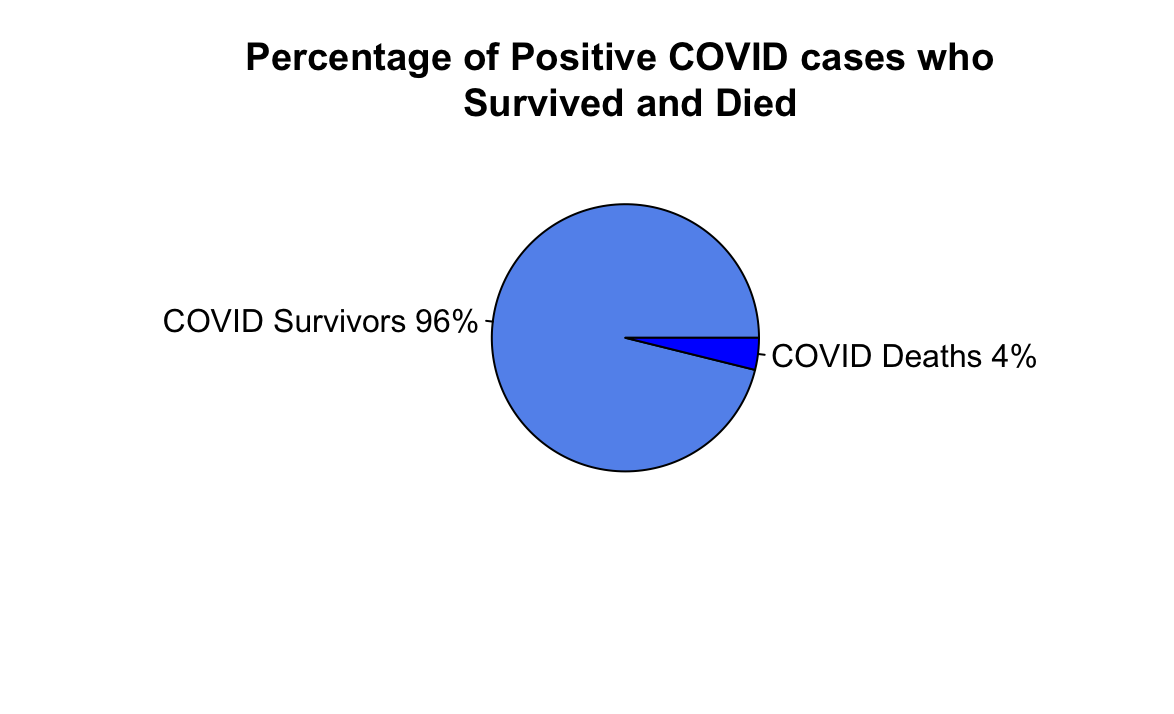
Hospitalization graph is a key indicator for understanding the severity of this disease and the pandemic’s impacts on the health care system.



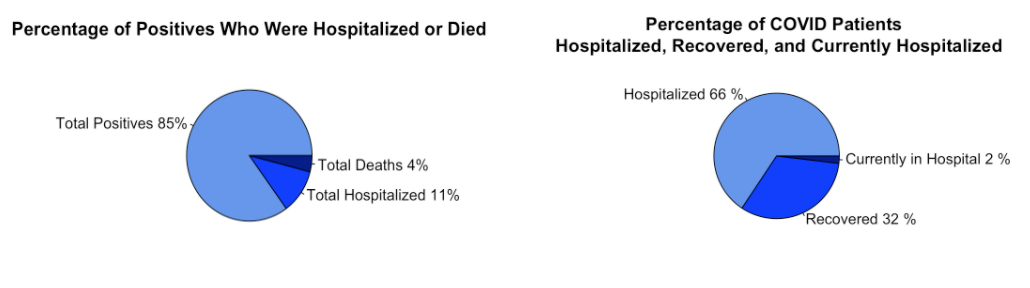
*The above graphs show a grid scatter plot of deaths confirmed and hospitalized against months.* Based on death data, the trend remains declining, and remains above the epidemic threshold. This number, of course, will always rise, but will also – eventually – plateau. A cumulative total can never fall.

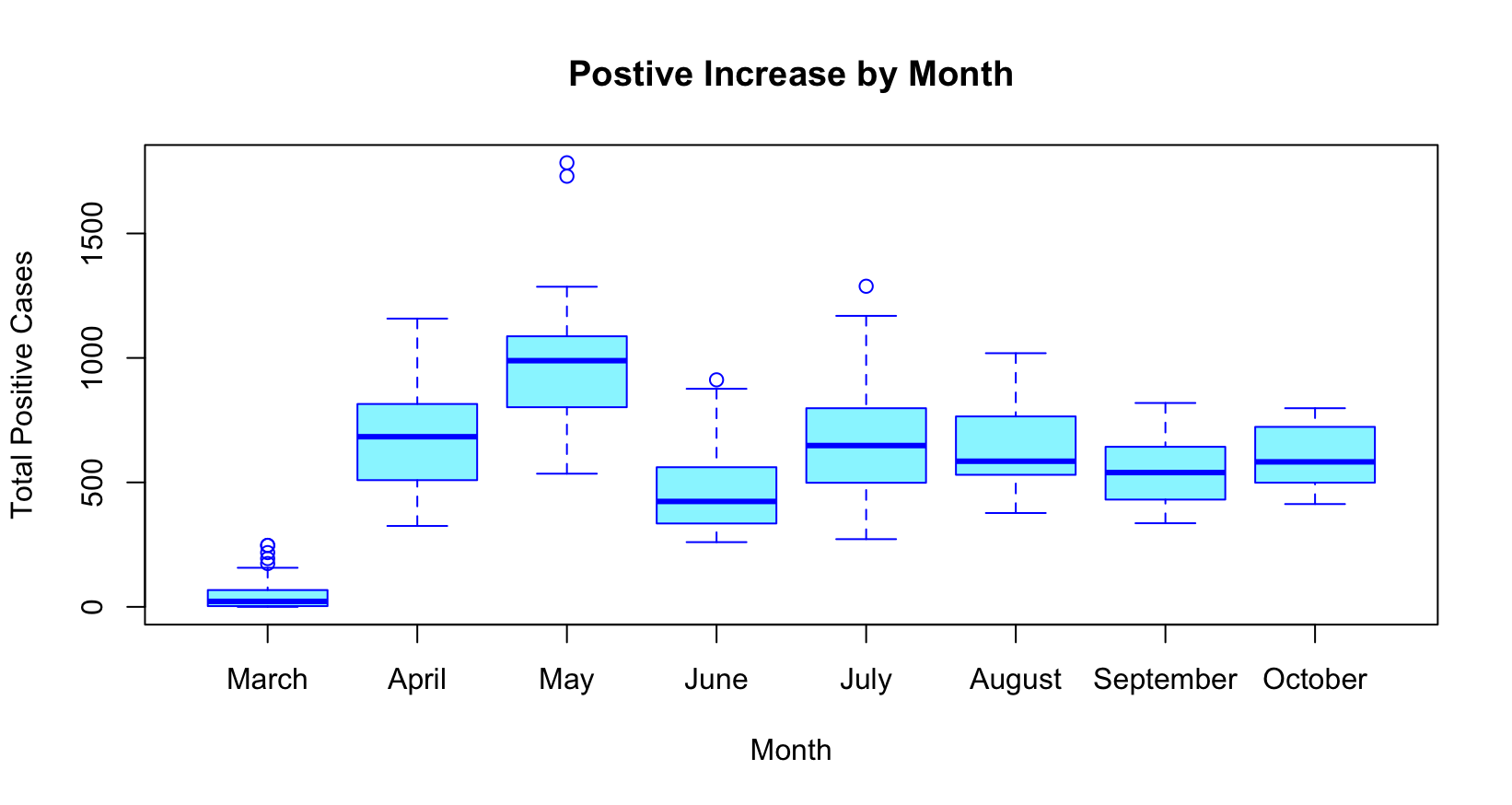
The steep upward moving line on the hospitalized graph indicates an increase in the number of hospitalised cases. The steeper the slope, the faster the total is increasing. Since the outbreak, overall weekly hospitalization rates have increased.

**d)**. A basic visualization of the total number of Marylanders who contacted COVID-19 and the number who died was deemed necessary for the layman’s understanding.



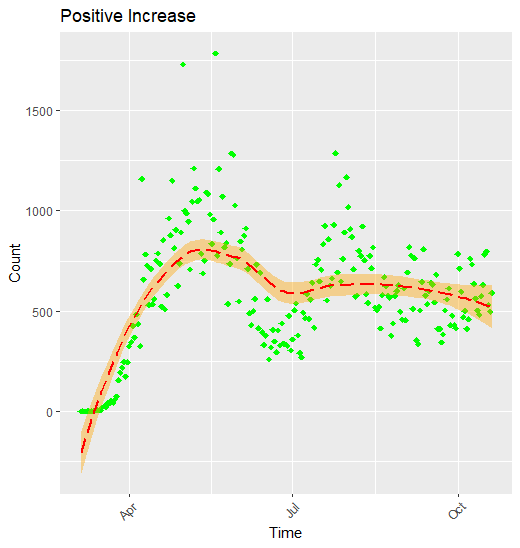
This pie chart illustrates that 4% of Marylanders who contracted COVID-19 died. Another way to discuss is the death-to-case ratio which is the number of COVID-19 deaths divided by the number of COVID-19 cases within a certain time interval - our dataset is from March 5, 2020 to October 20, 2020. We calculated this to be 38.7. However, the purpose of data visualization is to communicate information as clearly as possible. A death-to-case ratio might be meaningful to an epidemiologist, but this simple pie chart can communicate the severity of COVID-19 much more effectively in layman’s terms. If the entire Maryland population (6,046,000) were to become infected, 241,840 Marylanders would die.

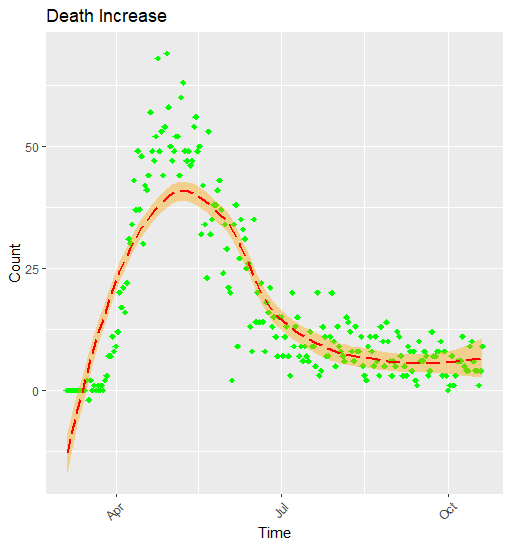


The above 2 pie charts contrast the differences in COVID-19 severity. While only 15% of patients are hospitalized (we are assuming that most deaths were hospitalized), once a patient enters the hospital their survival rate dramatically decreases. One of the most important aspects of lockdown measures is their ability to ameliorate the limitations of the state’s hospital capacity. This pie chart uses the cumulative numbers over the spring, summer, and fall to illustrate the recovery rate. The new discoveries that are being made regarding the management and treatment of novel COVID-19 probably impact the recovery rate, and thus are a major confounder to making associations between lockdown measures and recovery rates over time.While our ultimate goal is to do a time series analysis of the trends in positive cases, hospitalization, recoveries, and deaths to examine whether the lockdown orders had a measurable effect in Maryland. We used the dataset element “positiveIncrease'' instead of the “positive” because the latter is cumulative while the positive increase was a unique daily number. The box plot shows the month of May with a record high in the number of positive increases and March with the least. April and July seem alike with similar numbers of cases recorded, and June, August, September and October all share similar ranges of positive increase cases.

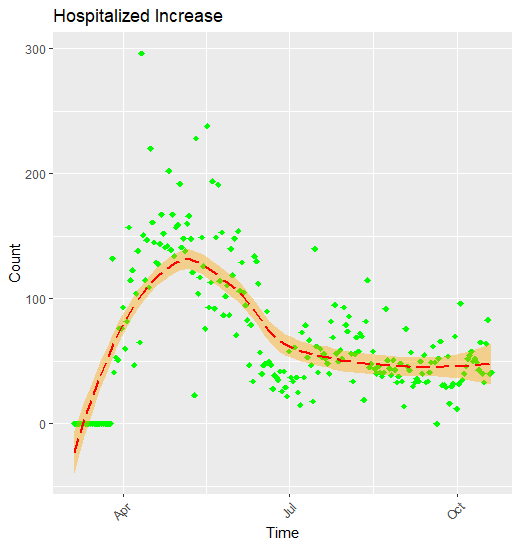
We are using the boxplot to brainstorm and to find clues. We know, for example, that there is a lag between when the orders are issued and when we can expect to see changes in the number of positive cases, yet we do not know how exactly long it takes for the changes to make an impact. On March 30th the governor issued a stay-at-home order, and on April 15th the face mask order was issued. We see a drop in numbers in June after 2 months of major shutdowns. The highest daily cases (around 1200) is in May. Yet, it could possibly be due to an increase in testing capacity. Stage 1 of reopening begins on May 13th, and we see the numbers rise again in July approximately 6 weeks after. The outlier in July could be as a result of the 4th of July celebrations. Schools reopened at the end of August, but the medians (around 500 cases a day) have stayed fairly close together thus far. We will examine these trends in further detail in our next deliverable.

Visualizing COVID-19 patterns lets people understand how and where the pandemic progresses more concretely. Many states are now gradually experimenting by relaxing standards to see whether COVID-19 can be evaluated.. This should give us an indication of whether the spread is easing or worsening. We get daily COVID-19 updates with lots of data, figures, and graphs to see if we flatten the curve with respect to the number of the new cases. The majority of these are based only on the total number of new cases confirmed or the daily number. This does not include ample data as to whether the situation improves, stabilizes or deteriorates. That's why we have to take into account the number of people tested daily for COVID-19 .

*Above is the scatterplot with loess regression line to show the trend of positive increase in cases against months.* According to our dataset **,** PositiveIncrease is the daily increase in positive, which measures Cases (confirmed plus probable) calculated based on the previous day’s value. We used the dataset element “positive Increase'' instead of the “positive” because the latter is cumulative while the positive increase was a unique daily number. If we observe the graph initially there is a striking upward trend, then started to decline this is because of initial phase 1 of reopening then again it started to incline due to reopening of schools and partial reopening of business.

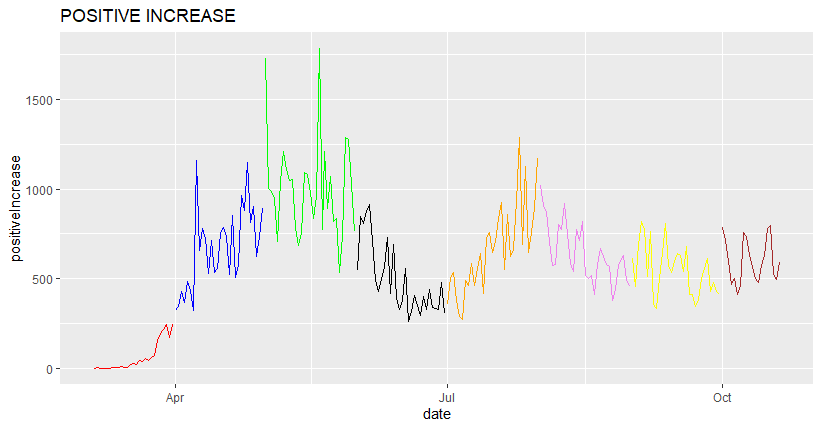


*Above is the scatterplot with loess regression line to show the trend of death increase in cases against months.* According to our dataset, deathIncrease is Daily increase in death, calculated from the previous day’s value. COVID-19 death rates are truly dropping as of mid-October. The elevation angle of the daily deaths curve gradually shifted from upward to downward pressure and tended to level out. The number of deaths caused by COVID-19 is one key metric that is often referred to, but as with other COVID metrics, it is a challenge to measure accurately. The issues involved in measuring COVID-19 deaths and argue that the change in the number of directly observed COVID-19 deaths is the most reliable and timely approach when using deaths to judge the trajectory of the pandemic in the Maryland, which is critical given the current inconsistencies in testing and limitations of hospitalization data.

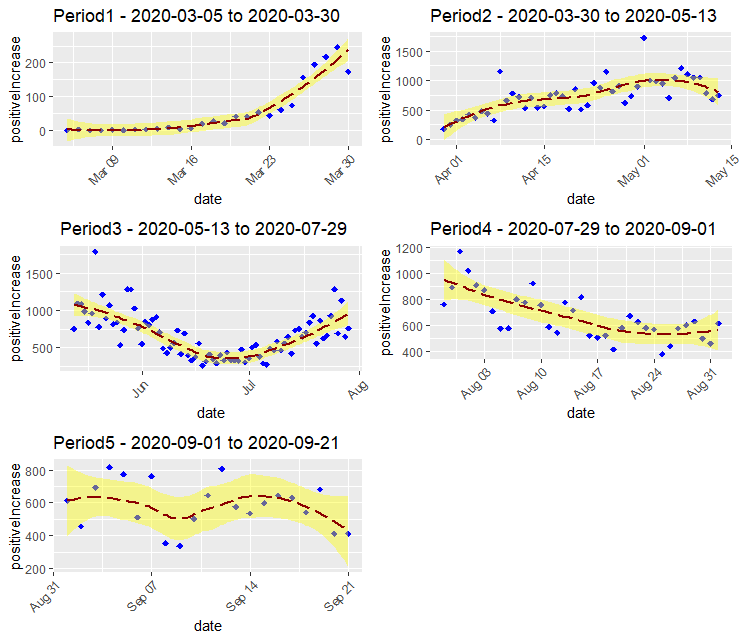


*Above is the scatterplot with loess regression line to show the trend of hospitalized increase in cases against months.* According to our dataset, hospitalizedIncrease Is Daily increase in hospitalizedCumulative, calculated from the previous day’s value. Hospitalization graph is a key indicator for understanding the severity of this disease and the pandemic’s impacts on the health care system. On a chart of hospitalized, look at how steeply the line is moving upward. The steeper the slope, the faster the total is increasing. Since the outbreak, overall weekly hospitalization rates have increased to an extent and then started to decline and maintained plateau as of mid-October data.

Without intervention steps, every 3-4 days is an exponential doubling of events. The number of cases in Maryland, unrestricted, would increase exponentially in the initial months. This would result in 1660 cases by the end of March, and 21742 cases by the end of April, if there were 3 cases on March 5. With this in mind, local authorities are taking drastic action. In order to deprive the virus of additional human targets, other authorities are enforcing preventive measures such as closing down schools and restricting social gatherings. As of May 19, Maryland had the most cases confirmed on a regular basis.



We grouped each month into a separate data frame and plotted line graph to see if months had different trends. Here the line graph shows the new daily cases from mid-March until October. Each month's trend is depicted with different color on the graph for easier analysis. In the daily number of confirmed cases we see high jumps and large fluctuations going back and forth. From the daily new cases data, it looks like there is a strongly decreasing trend in the number of confirmed cases in June.



As a part of stratified sampling method, we have decided to group dates into 5 periods for better understanding of trends.

Period1 indicates timeline from declaring a state of emergency to issuing stay-home order.

Period2 indicates timeline from issuing stay-home order to stage one of reopening.

Period3 indicates timeline from stage one of reopening to expansion of mask order.

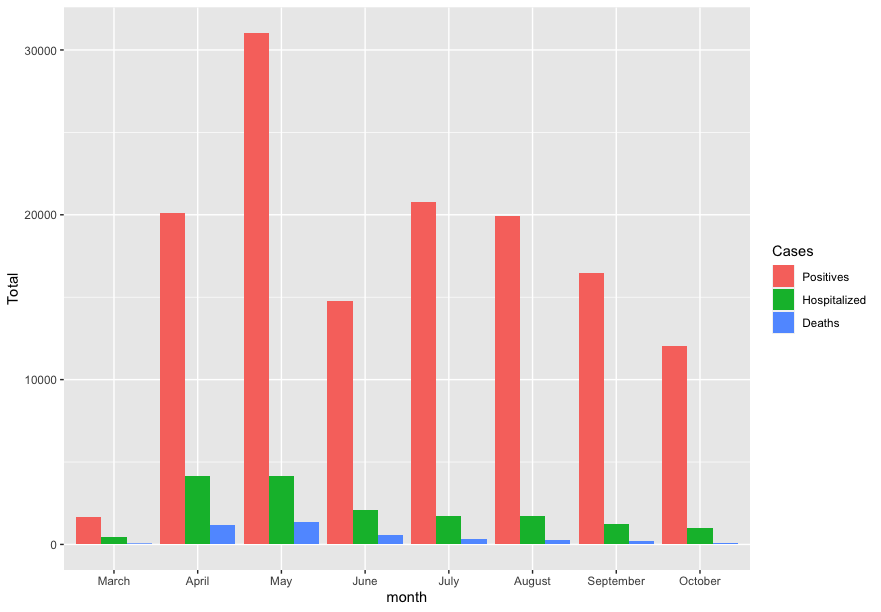
Period4 indicates timeline from expansion of mask order to stage 3 recovery plan.

Period5 indicates timeline from Stage 3 recovery plan to expansion of restaurants indoor capacity.

The graph shown above is grid chart showing scatterplots with regression line for Period1, Period2, Period3, Period4, and Period5

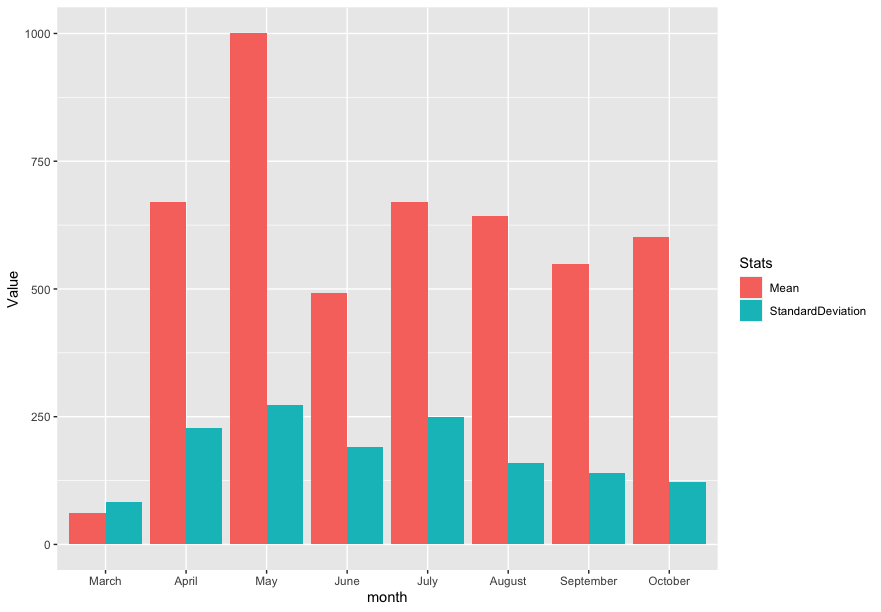
Period1 graph shows a sudden spike in trend as in this time period government mandates are in the initial stage. Period2 graph shows fluctuations but gradually increasing trend in positive cases. period3 graph shows a declining trend to some extent and then increasing this is due to the stage one of reopening. Soon after reopening cases came to control then suddenly a number of cases increased. Period4 graph shows declining trend throughout and plateau for some period. Period5 graph shows an increasing trend because of school reopening , positive cases in children increased this may be the potential reason and then declined due to stage 3 of the recovery plan.

*Grouped bar chart of sum of each positives, hospitalization, and deaths by month*



The above bar chart looks closer at the chosen categories, or months. We see that hospitalizations and deaths mimic positive cases rates. Positive cases are more likely to be impacted by mandates while hospitalizations and deaths are more likely to be related by hospital capacity and advancing medical knowledge about COVID-19.

*Grouped bar chart of mean and standard deviation of positive cases by month*



Thus, we chose to look closer at the means and standard deviations of the positive cases. However, we have not found any notable observations regarding each month’s data variance. The next step after visualization is to commence a statistical approach.

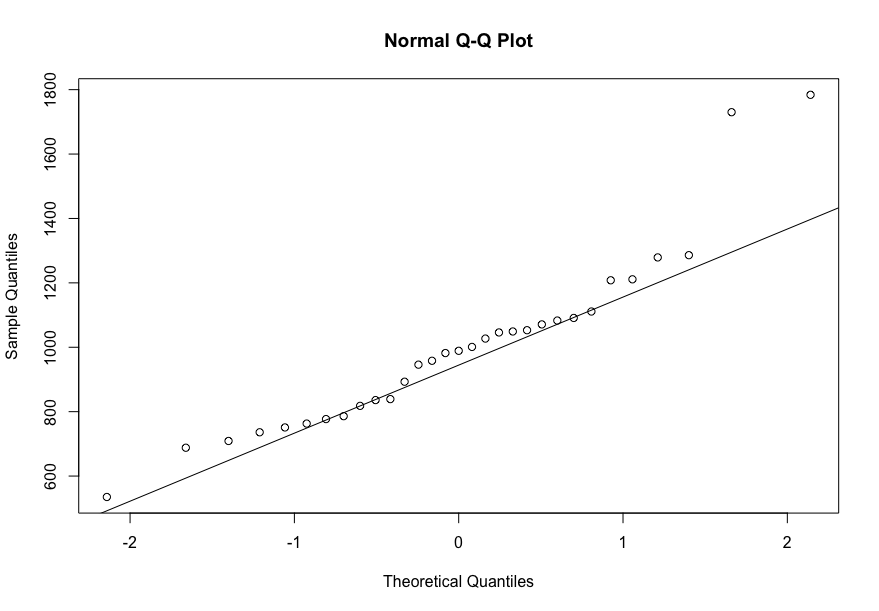
**f)** Analysis of Variance

We ran one-way ANOVA to test if there is statistical significance between months. However, before, we tested the data if it met the ANOVA model. We examined whether the assumptions of the data were upheld.

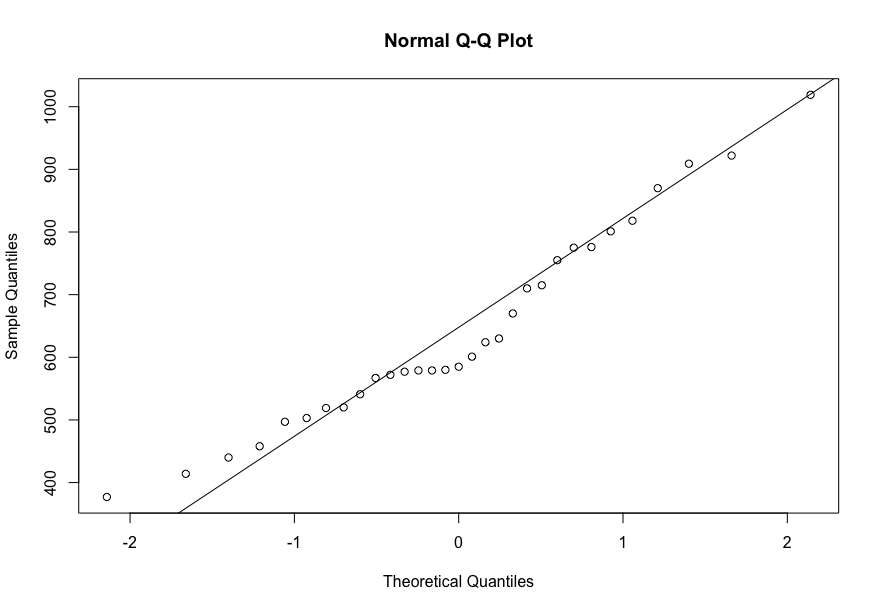
i) Assumption*: Normal Distribution of Each Group*

Below is a sample of the plots. We see some outliers in May.

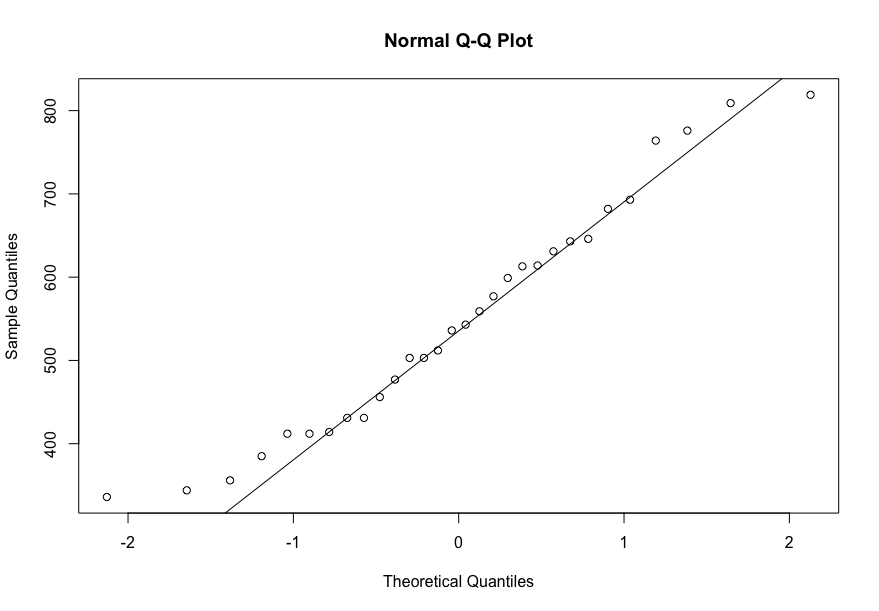
*For May:*



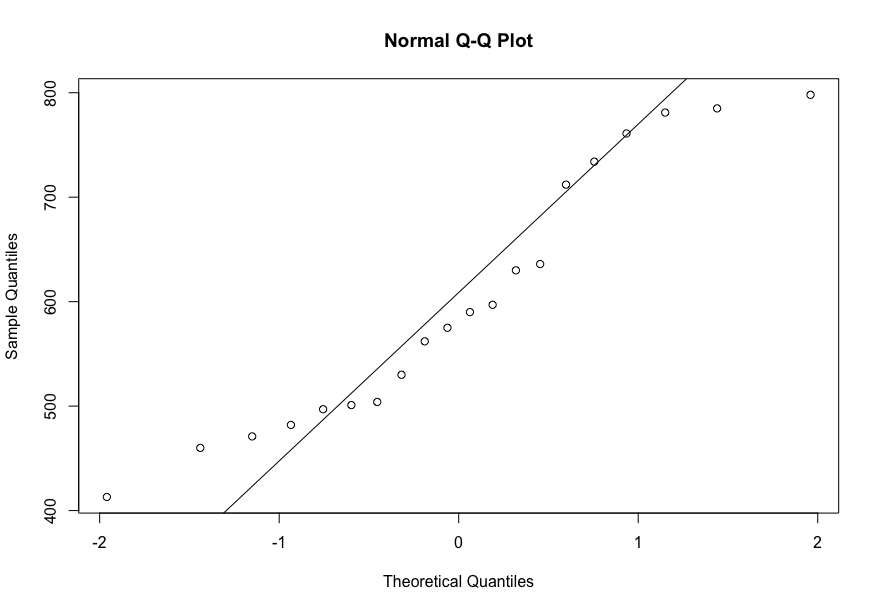
*For August:*



*For September:*

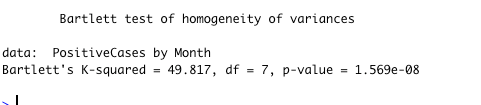


*For October:*

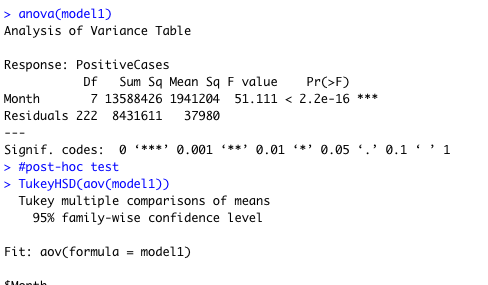


ii) Assumption: Homogeneity of Variances

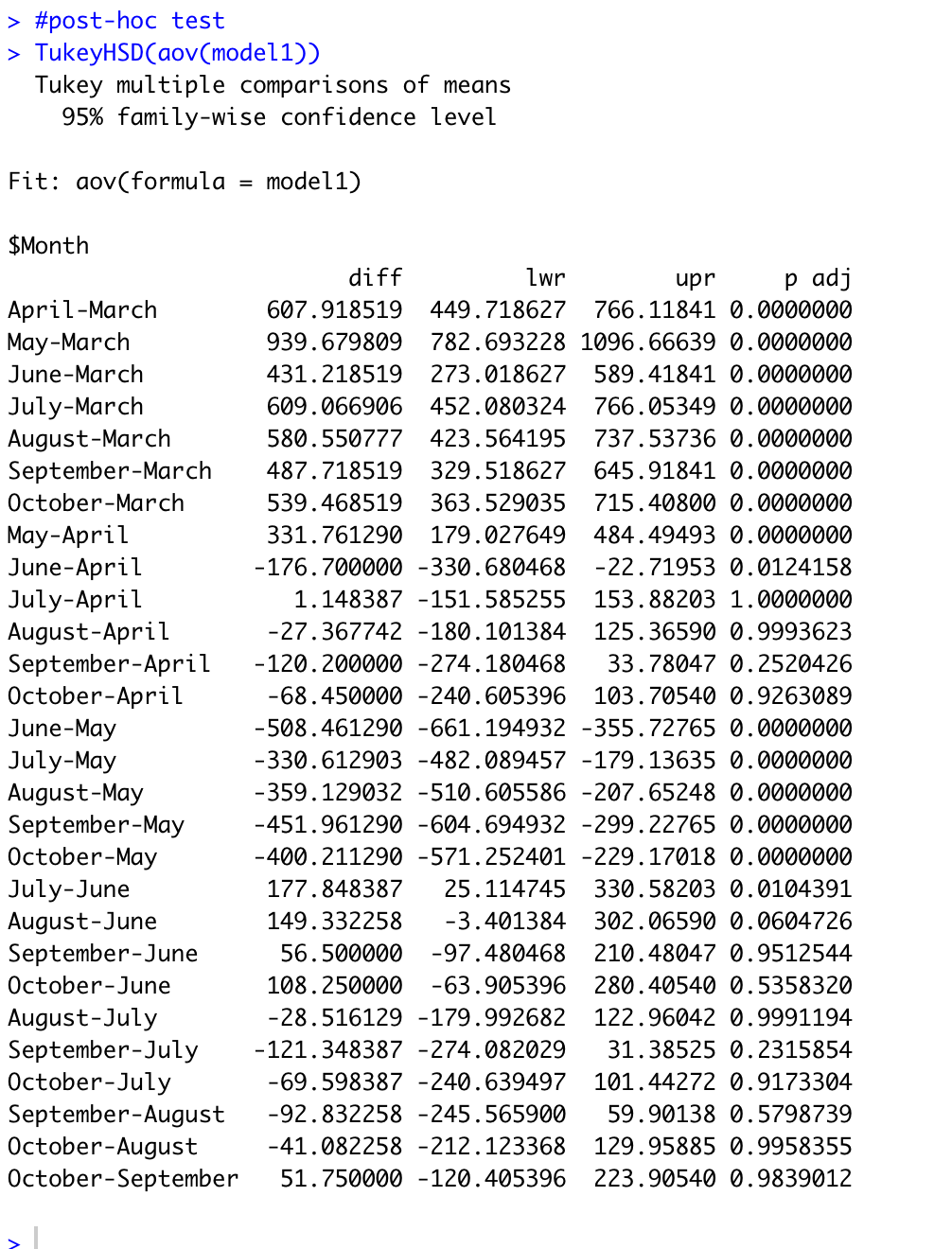
To test the assumption of homogeneity of variance in our dataset, we used the Bartlett Test which was statistically significant (p<0.01), that is this assumption appears to be upheld.



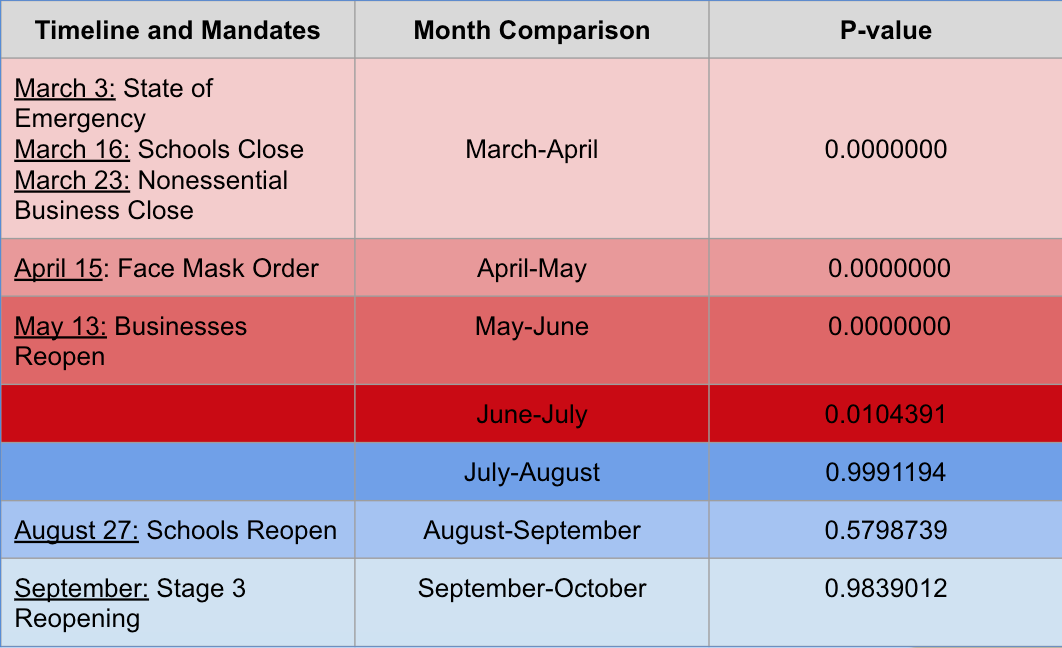
Finally, we ran the ANOVA model and found that there was a statistical significance difference.



The confidence interval we’ve chosen is 99%. The p-value is 2.2 x 10-6 i.e. than <0.01. Since this was statistically significant, a Tukey Test was used to conduct a post-hoc test in order to further examine the difference between the months. The R results follow:



We noted that May, the highest months, is statistically significant from all other months. June, the lowest month, is statistically significantly different from all other months, except for September and October. We chose to focus on successive months and summarized the results in the following table:



There is statistical significance between successive months from March until June. While a full shutdown was recommended, the mandates in April and May were still fairly conservative. However, the lockdowns eventually resulted in a June downward dip that was statistically significant. In July, the cases spiked again. From our previous, line graphs we can see the peak number of cases in July is mid-July which is 10 months after July 4. From July to October, the shifts in positive cases are not significant. The data suggests the plateau is a result of no major mandate orders were made during this period, thus the numbers stabilized.

**Discussion and Limitations**

The literature used event study analysis which enables them to test whether a single event, such as a mandate enforcement date, has a significant impact. A time series analysis seemed more appropriate but our data did not meet the assumptions of this model. Chiefly, it did have stationarity; the mean, variance, and autocorrelation did not stay consistent over time. Instead, we decided to take inspiration from the stratified sampling model and categorized time by month. Grouping the data into various months, made it convenient to manage given that our data set was incremental on a monthly basis.

The p-values may have been unusually low due to the fact that not all one-way ANOVA assumptions were met. Our interpretation has been conservative due to the fact that we did not use the most appropriate model for our data and hypotheses.

Moreover, we were not able to take into account some of the major contributing factors to COVID-19 infection rates like the enforcement of mandates and how well the population adhered to the social distancing portions of the mandate.

Our data was also not further stratified by age, health status, profession, or geographic area which may have facilitated a more narrow analysis of the impact of COVID-19.

In conclusion, our data suggests that the statistical difference between months can be correlated with mandates as well as public behavior.

***Appendix A***

***R code available from:***

[*https://drive.google.com/file/d/1p7Ri0x-NQt16qi0uASajXjJK7eY-ets1/view?usp=sharing*](https://drive.google.com/file/d/1p7Ri0x-NQt16qi0uASajXjJK7eY-ets1/view?usp=sharing)

***Dataset used available from:***

[*https://drive.google.com/file/d/1hIfuWDDOr4QD3QrZDdR6RGmWD8mFuU8M/view?usp=sharing*](https://drive.google.com/file/d/1hIfuWDDOr4QD3QrZDdR6RGmWD8mFuU8M/view?usp=sharing)

***Appendix B***

***Cases***

**Cases (confirmed plus probable) - API field name: positive**

Total number of confirmed plus probable cases of COVID-19 reported.

**New cases - API field name: positiveIncrease**

The daily increase in API field positive, which measures Cases (confirmed plus probable) calculated based on the previous day’s value.

**Probable Cases - API field name: probableCases**

Total number of probable cases of COVID-19 as reported by the state. A probable case is someone who tests positive via antigen without a positive PCR or other approved nucleic acid amplification test (NAAT), someone with clinical and epidemiological evidence of COVID-19 infection with no confirmatory laboratory testing performed for SARS-CoV-2, or someone with COVID-19 listed on their death certificate with no confirmatory laboratory testing performed for SARS-CoV-2.

***PCR tests***

**Confirmed Cases or Positive PCR tests (people) - API field name: positiveCasesViral**

Total number of unique people with a positive PCR or other approved nucleic acid amplification test (NAAT), as reported by the state or territory. This is equivalent to a confirmed case.

**Negative PCR tests (people) - API field name: negative**

Total number of unique people with a completed PCR test that returns negative.

**Negative PCR tests (specimens) - API field name: negativeTestsViral**

Total number of completed PCR tests (or specimens tested) that return negative as reported by the state.

**Pending - API field name: pending**

Total number of viral tests that have not been completed as reported by the state.

**Positive PCR tests (specimens) - API field name: positiveTestsViral**

Total number of completed PCR tests (or specimens tested) that return positive as reported by the state.

**Total PCR tests (people ) - API field name: totalTestsPeopleViral**

Total number of unique people tested at least once via PCR testing, as reported by the state. The count for this metric is incremented up only the first time an individual person is tested and their result is reported. Future tests of the same person will not be added to this count.

**Total PCR tests (specimens) - API field name: totalTestsViral**

Total number of PCR tests (or specimens tested) as reported by the state or territory. The count for this metric is incremented up each time a specimen is tested and the result is reported.

**Total PCR tests (test encounters) - API field name: totalTestEncountersViral**

Total number of people tested per day via PCR testing as reported by the state. The count for this metric is incremented up by one for each day on which an individual person is tested, no matter how many specimens are collected from that person on that day. If an individual person is tested twice a day on three different days, this count will increment up by three.

***Antibody tests***

**Negative antibody tests (people) -API field name: negativeTestsPeopleAntibody**

The total number of unique people with completed antibody tests that return negative as reported by the state.

**Negative antibody tests (specimens) - API field name: negativeTestsAntibody**

The total number of completed antibody tests that return negative as reported by the state.

**Positive antibody tests (people) - API field name: positiveTestsPeopleAntibody**

The total number of unique people with completed antibody tests that return positive as reported by the state .

**Positive antibody tests (specimens) - API field name: positiveTestsAntibody**

Total number of completed antibody tests that return positive as reported by the state.

**Total antibody tests (people) - API field name: totalTestsPeopleAntibody**

The total number of unique people who have been tested at least once via antibody testing as reported by the state.

**Total antibody tests (specimens) - API field name: totalTestsAntibody**

Total number of completed antibody tests as reported by the state.

***Antigen tests***

**Positive antigen tests (people) - API field name: positiveTestsPeopleAntigen**

Total number of unique people with a completed antigen test that returned positive as reported by the state.

**Positive antigen tests (specimens) - API field name: positiveTestsAntigen**

Total number of completed antigen tests that return positive as reported by the state.

**Total antigen tests (people) - API field name: totalTestsPeopleAntigen**

Total number of unique people who have been tested at least once via antigen testing, as reported by the state.

**Total antigen tests (specimens) - API field name: totalTestsAntigen**

Total number of completed antigen tests, as reported by the state.

***Hospitalization***

**Cumulative hospitalized/Ever hospitalized - API field name: hospitalizedCumulative**

Total number of individuals who have ever been hospitalized with COVID-19.

**Cumulative in ICU/Ever in ICU - API field name: inIcuCumulative**

Total number of individuals who have ever been hospitalized in the Intensive Care Unit with COVID-19.

**Cumulative on ventilator/Ever on ventilator - API field name: onVentilatorCumulative**

Total number of individuals who have ever been hospitalized under advanced ventilation with COVID-19.

**Currently hospitalized/Now hospitalized - API field name: hospitalizedCurrently**

Individuals who are currently hospitalized with COVID-19.

**Currently in ICU/Now in ICU - API field name: inIcuCurrently**

Individuals who are currently hospitalized in the Intensive Care Unit with COVID-19.

**Currently on ventilator/Now on ventilator - API field name: onVentilatorCurrently**

Individuals who are currently hospitalized under advanced ventilation with COVID-19.

**New total hospitalizations - API field name: hospitalizedIncrease**.

Daily increase in hospitalizedCumulative, calculated from the previous day’s value.

***Outcomes***

**Deaths (confirmed and probable) - API field name: death.**

Total fatalities with confirmed OR probable COVID-19 case diagnosis.

**Deaths (confirmed) - API field name: deathConfirmed.**

Total fatalities with confirmed COVID-19 case diagnosis

**Deaths (probable) - API field name: deathProbable.**

Total fatalities with probable COVID-19 case diagnosis

**New deaths - API field name: deathIncrease.**

Daily increase in death, calculated from the previous day’s value.

**Recovered - API field name: recovered.**

Total number of people that are identified as recovered from COVID-19.

**Date - API field name: date**

Date on which data was collected by The COVID Tracking Project.

**References**

1. CDC. (2020, August 7). Coronavirus Disease 2019 (C)VID-19). Centers for Disease control and Prevention.

<https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/cloth-face-cover-guidance.html>

1. Coronavirus disease 2019. (2020, March 3). Wikipedia.

[https://en.wikipedia.org/wiki/Coronavirus\_disease\_2019‌](https://en.wikipedia.org/wiki/Coronavirus_disease_2019%E2%80%8C)

1. *Dr. Fauci Sets the Record Straight About Mask. (n.d). Www. Msn.Com.* Retrieved on October 27, 2020. From <https://www.msn.com/en-us/health/medical/dr-fauci-sets-the-record-straight-about-masks/ar-BB19Aof3>
2. *Estimating mortality from COVID-19*. (n.d.). Www.Who.Int. <https://www.who.int/news-room/commentaries/detail/estimating-mortality-from-covid-19>
3. *Hypothesis Testing - Examples and case studies. (n.d).*

<https://www2.stat.duke.edu/courses/Fall11/sta10/STA10lecture21.pdf>

1. Jacobs P, Ohinmaa AP. The enforcement of statewide mask wearing mandates to prevent COVID-19 in the US: an overview. F1000Res. 2020;9:1100.