Predict 2020 USA Presidential Election COVID-19 Correlation

Saloni Patel
Stanford Online High School, Redwood City, CA 94063
saloni05@ohs.stanford.edu

Mason Chen
Stanford Online High School, Redwood City, CA 94063
mason05@ohs.stanford.edu

Patrick Giuliano
Abbott Structural Heart, Menlo Park, CA 94025
patrick.giuliano@gmail.com

Abstract

This project investigates different strategies Trump can utilize in reopening states during the COVID-19 pandemic. During the pandemic, the President and state governors face the challenge of deciding when to reopen the states and whether they should all be opened all at once or based on their individual situations. Based on several calculated variables, such as past win margins of swing states, infected cases, deaths, and unemployment increases for 16 different swing states from past elections, the authors draw conclusions on which states President Trump should be put into consideration to "liberate" or "reopen" to not only safely reopen, but to maximize his chances of winning the 2020 election. With the calculated and collected variables, a statistical model is created to aid in decision making. Although the safest option is to stay closed, many state economies and the overall national economy suffer due to the closure. Trump's pro-economy campaign must wisely select which states to liberate based not only on unemployment rates, but chances of winning that state in the upcoming 2020 election. This project plays special attention to Michigan, Minnesota, and Virginia, which were called out in President Trump's tweet on April 17, 2020. Ultimately, this project concludes that Minnesota is a safe state to liberate, Michigan is too risky, and Virginia can be liberated, but the authors advise against it. Additionally, two possible scenarios are also covered in this paper. This project demonstrates a practical and statistical modeling framework considering social and political science factors.

Keywords
U.S. Presidential Election, COVID-19, Predictive Modeling

1. Introduction

On April 17, 2020 President Trump had urged Minnesota, Michigan and Virginia to "liberate" themselves in an apparent criticism of stay-at-home orders. If the states were to "liberate" they would no longer have to follow stay-at-home-orders that were put in place in an attempt to impede the spread of the virus. President Trump's tweets had signaled support for protestors who were demanding that the economic-crippling stay-at-home measures be eased, or removed. Of course, relenting on the social-distancing orders risks a resurgence of infections and deaths, but what about effects on business? On April 22, President Trump said he "strongly" disagreed with the move to reopen Georgia, which contradicted his previous statements on "liberating" the other states and a source who said he agreed with it. President Trump must address why he chose to "liberate" or not "liberate" some states. This project explores and predicts what President Trump's decisions should be on liberating. First, we establish a statistical model to study COVID-19 effects and Jobless Rate on the 2020 US Presidential election result on swing states. In order to control the spread of COVID, President Trump must consider that if he opens too late, the economy will be in trouble (which would cause him to lose votes), and if he opens too early then there will be an increased risk of contracting the virus (which would also lose him votes). The authors explore what are the risks and consequences of liberating Michigan, Minnesota, and Virginia, and which swing states should President Trump reconsider reopening for business.

2. The Model

To begin building the model, the authors first establish a statistical model using infection cases, deaths, and unemployment rates. Next, the risks and consequences of liberating Michigan, Minnesota, and Virginia are analyzed, along with two possible scenarios that President Trump should consider reopening.

2.1 Background Research

COVID-19 is a serious, novel virus causing global health issues ranging from headaches to coughs and shortness of breath. COVID-19 has spread to more 180 countries, and the United States leads in most cases with more than 6 million. COVID-19 has hindered global events and caused trouble across the world. Additionally, unemployment rates in the United States reached a record high of 14.7 percent in April 2020 due to shelter-in-place and business closures to contain the spread of the virus.

The proliferation of COVID-19 could have significant and lasting effects on the policy, economy, and infrastructure of the United states, where the fate of the Presidential Election rests in the hands of a small body of electorates: the electoral college. The electoral college was established by the United States constitution in each state based on the population of the state, as seen in the diagram. In every election, whichever president wins the most electoral votes in a state wins that entire state. Figure 1 illustrates the number of electoral colleges in each state. The number of electoral college votes will be a variable in the statistical model.

Additionally, the fate of the Presidential Election also depends on several battleground or "swing" states. Swing states are states that are easily "swung" by either Democrats or Republicans because they usually do not have a fixed, large population of either Democrats or Republicans. In Figure 2, the lighter shades of blue and red are swing states. Swing states are usually important because they can be won by either party, so each of the candidates from each side focus heavily on the swing states in attempts to sway them to their side. In this project 16 swing states from the two most recent elections are analyzed to determine whether President Trump should consider advocating for reopening them. It is important to remember that President Trump does not have legal authority to actually reopen the states, but he can only recommend governors to consider reopening their states.

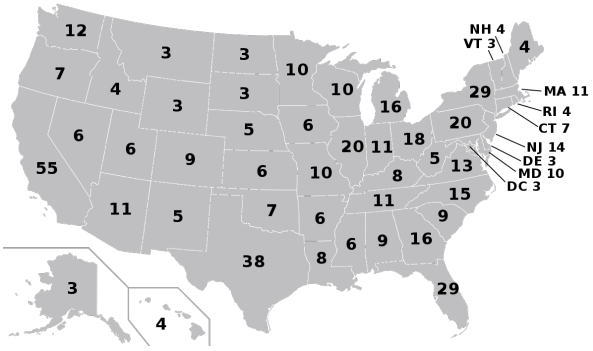

Figure 1. Electoral Colleges

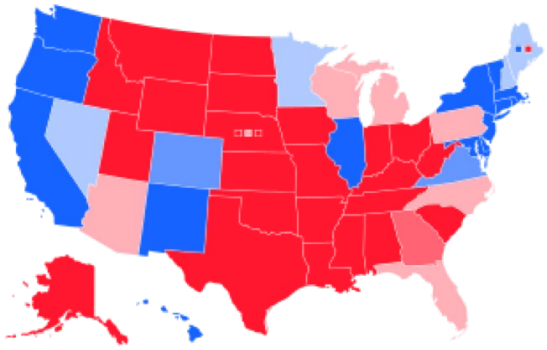

Figure 2. Swing States

2.2 Hypothesis

Due to the rather unstable economic conditions and high unemployment rates as well as increasing infected cases and deaths caused by the virus in the United States, it was hypothesized that President Trump should not liberate any states unless unemployment rates are in serious trouble.

2.3 Model Building

To set-up the model, the COVID-19 infected cases and death rates per 100k statistics were uploaded. The authors also added annual unemployment increase rates for each of the swing states. These three variables were chosen because the number of infected cases indicates the spread of the virus and the number of death rates indicates the spread as well as the strength of the healthcare system in each state. The unemployment increase rates indicate the effect COVID-19 has taken and will take on the businesses and overall economy of that state. Table 1 displays the infected cases per 100k, death number per 100k and annual unemployment increase for 16 swing states.

State	Infected Cases per 100k	Death Number per 100k	Annual Unemployment Increase
Oregon	45	2	4.63%
New Mexico	88	3	4.42%
Colorado	169	7	1.98%
Virginia	100	3	3.03%
Nevada	121	5	8.18%
Minnesota	42	2	6.74%
Michigan	315	24	6.85%
New Hampshire	41	3	7.34%
Wisconsin	75	4	5.96%
Pennsylvania	257	10	7.73%
Florida	123	4	0.76%
North Carolina	63	2	5.16%
Arizona	68	3	1.52%
Ohio	99	4	7.31%
Iowa	92	2	4.44%
Georgia	172	6	7.69%

Table 1. Identified Infection Cases, Deaths, and Unemployment Rate in Swing States

3. Data & Statistical Analysis

In order to create a composite winning margin for each swing state from the past, the Z-Standardization is applied to z-scores (z-infected, z-death, z-unemployment) (see Table 3) with different distribution so they can still be compared equally, and the liberate indices are derived to build the model and make predictions.

3.1 Composite Winning Margin Calculations

16 swing states were identified based on the past two presidential elections and the winning margins for those swing states were identified as well as the number of electoral college votes in each state. Only swing states are applied in this model because swing states tend to be swayed more easily and are often targeted in election campaigns. Win margins of the 16 swing states from the past two elections, 2016 and 2012, were used because President Trump participated in the 2016 election and Joe Biden was a vice presidential candidate in the 2012 election. The authors gave the 2016 result twice the weight because it was more recent than the 2012 election. Another two times the weight was added for 2016 because President Trump was present in the 2016 election running as president, while Joe Biden was present in 2012 as a candidate for vice president. In total the 2016 result will have four times the weight compared to 2012. The 2016 result is multiplied by 4, the 2012 result is added, and divided by 5 to derive the composite winning margin by averaging, which can be seen in Equation 1. If the winning margin is negative, President Trump loses, and if it is positive, President Trump wins. Table 2 depicts how Michigan had voted red, but is predicted to vote blue this year as well as Pennsylvania, both large swing states.

{[(“2016 Result") * 4] +("2012 Result")} / 5 = **Composite Winning Margin**

Equation 1. Composite Winning Margin

State	2020 Votes	2016 Result	2012 Result	2016-2012 Composite
Oregon	7	-10.98%	-12.08%	-11.20%
New Mexico	5	-8.22%	-10.15%	-8.61%
Colorado	9	-4.91%	-5.36%	-5.00%
Virginia	13	-5.32%	-3.88%	-5.03%
Nevada	6	-2.42%	-6.68%	-3.27%
Minnesota	10	-1.52%	-7.69%	-2.75%
Michigan	**16**	0.23%	-9.50%	-1.72%
New Hampshire	4	-0.37%	-5.58%	-1.41%
Wisconsin	10	0.77%	-6.94%	-0.77%
Pennsylvania	**20**	0.72%	-5.38%	-0.50%
Florida	29	1.20%	-0.88%	0.78%
North Carolina	15	3.66%	2.04%	3.34%
Arizona	11	3.50%	9.03%	4.61%
Ohio	18	8.13%	-2.98%	5.91%
Iowa	6	9.41%	-5.81%	6.37%
Georgia	16	5.09%	7.82%	5.64%

Table 2. 2016-2012 Composite Win Margin of Swing States

3.2 Z-Standardization, Z-COVID Index, Z-Risk

The Z-Standardization is a process that will compare the z-scores of samples with different distributions. It will convert the separate distributions of the z-scores of the infected, death, and unemployment rates into a standardized distribution, so the z-scores can be compared equally. The authors use this statistical tool to avoid any sampling mean and variance bias. The z-score represents the number of standard deviations the sample is from the center mean. A positive z-score indicates that the sample is above average, and negative is below average. To get the z-scores, the authors subtracted the value by the sample mean and then divided by the sample standard deviation. The three z-scores areh Z-Infected, Z-Death, and Z-Unemployment. In Table 3, Michigan has an above average rate of infection by two standard deviations and an above average death rate by more than 3 standard deviations. Oregon, New Mexico, and Arizona are below average for all infected cases, deaths, and unemployment.

(Infected Cases per 100k - 151.7) / 77.6 = **Z-Infected score**
Equation 2. Z-Infected score

(Death Number per 100k - 6.01) / 5.45 = **Z-Death score**
Equation 3. Z-Death score

(Annual Unemployment Increase - 0.054) / 0.0239 = **Z-Unemployment score**
Equation 4. Z-Unemployment score

Next the authors derived the Z-COVID index by averaging out the previous three z-scores (Z-Infected, Z-Death, and Z-Unemployment). Now with the Z-COVID index, 2016-2012 composite winning margin, and the standard deviation of the 2016-2012 average winning margins (which is simply the winning margin of 2012 and 2016 added and then divided by 2 to get the average), the predicted 2020 win margin for each swing state is calculated. These three values are z-scores and once averaged out they produce the 2020 win margin, which accounts for all those variables. The standard deviation for the average 2016-2012 is about 0.0518. Once again a positive 2020 win margin is favorable for President Trump, and a negative margin would mean Trump loses in that state.

State	Z-Infected	Z-Death	Z-Unemployment
Oregon	-1.38	-0.74	-0.32
New Mexico	-0.82	-0.55	-0.41
Colorado	0.22	0.18	-1.43
Virginia	-0.67	-0.55	-0.99
Nevada	-0.40	-0.19	1.16
Minnesota	-1.41	-0.74	0.56
Michigan	2.10	3.30	0.61
New Hampshire	-1.43	-0.55	0.81
Wisconsin	-0.99	-0.37	0.23
Pennsylvania	1.36	0.73	0.97
Florida	-0.37	-0.37	-1.94
North Carolina	-1.14	-0.74	-0.10
Arizona	-1.08	-0.55	-1.62
Ohio	-0.68	-0.37	0.80
Iowa	-0.77	-0.74	-0.40
Georgia	0.26	0.00	0.96

Table 3. Z-Infected, Z-Death, Z-Unemployment

[("Z-Infected") +("Z-Death") + ("Z-Unemployment")] / 3 = **Z-COVID Index**
Equation 5. Z-COVID Index

("2016-2012 AVG") -("Z-COVID") * 0.0518 = **2020 Win Margin**
Equation 6. 2020 Win Margin

State	2020 Votes	2016 Result	2012 Result	2016-2012 AVG	Z-Infected	Z-Death	Z-Unemployment	Z-COVID	2020 Win Margin
Oregon	7	-10.98%	-12.08%	-11.20%	-1.38	-0.74	-0.32	-0.81	-7.00%
New Mexico	5	-8.22%	-10.15%	-8.61%	-0.82	-0.55	-0.41	-0.59	-5.53%
Colorado	9	-4.91%	-5.36%	-5.00%	0.22	0.18	-1.43	-0.34	-3.23%
Virginia	13	-5.32%	-3.88%	-5.03%	-0.67	-0.55	-0.99	-0.74	-1.22%
Neveda	6	-2.42%	-6.68%	-3.27%	-0.40	-0.19	1.16	0.19	-4.28%
Minnesota	10	-1.52%	-7.69%	-2.75%	-1.41	-0.74	0.56	-0.53	-0.01%
Michigan	16	0.23%	-9.50%	-1.72%	2.10	3.30	0.61	2.00	-12.10%
New Hampshire	4	-0.37%	-5.58%	-1.41%	-1.43	-0.55	0.81	-0.39	0.60%
Wisconsin	10	0.77%	-6.94%	-0.77%	-0.99	-0.37	0.23	-0.37	1.17%
Pennsylvania	20	0.72%	-5.38%	-0.50%	1.36	0.73	0.97	1.02	-5.79%
Florida	29	1.20%	-0.88%	0.78%	-0.37	-0.37	-1.94	-0.89	5.41%
North Carolina	15	3.66%	2.04%	3.34%	-1.14	-0.74	-0.10	-0.66	6.75%
Arizona	11	3.50%	9.03%	4.61%	-1.08	-0.55	-1.62	-1.08	10.23%
Ohio	18	8.13%	-2.98%	5.91%	-0.68	-0.37	0.80	-0.08	6.34%
Iowa	6	9.41%	-5.81%	6.37%	-0.77	-0.74	-0.40	-0.64	9.66%
Georgia	16	5.09%	7.82%	5.64%	0.26	0.00	0.96	0.41	3.53%

Table 4. Z-COVID Index and 2020 Win Margin

Now the authors derive the Z-Risk 2 by 1 and 1 by 1 indexes. The Z-Risks assess the risk of liberating each state. To do this the previous three z-scores (Z-Infected, Z-Death, and Z-Unemployment) were averaged to get the Z-risk 2 by 1. Some may think that this is biased because COVID-19 appears to get more importance, since there are two values for COVID-19 and only one for unemployment. Thus, the authors decided to calculate the Z-Risk 1 by 1, so COVID and unemployment rates are assessed equally. The formula for the z-risk 1 by 1 is used (see below).

("Z-Infected") + [("Z-Death") -("Z-Unemploy")] = **Z-Risk 2-1**
Equation 7. Z-Risk 2-1

{[("Z-Infected") + ("Z-Death")] / 2} - ("Z-Unemploy") = **Z-Risk 1-1**
Equation 8. Z-Risk 1-1

Next, the z-risk 1 by 1 is used to calculate the Liberate Index by multiplying the z risk 1 by 1 with the 2020 votes or the electoral college votes. The Liberate Index represents the chance of the state switching sides from red to blue or blue to red. If the Liberate Index is far from zero, then the chance of the state switching sides is low. If the Liberate Index is or close to zero, then the chance of the state switching sides is high. The states in red are hard for President Trump to get back, since they are not likely to change, while the states in green are already with President Trump since they are also not likely to change. The states in white are likely to change, since their liberate index is low. These states in white would now be classified as true swing states, since they have a higher chance of switching sides compared to the swing states in red and green.

("2020 Votes") *("Z-Risk 1-1") = **Liberate Index**
Equation 9. Liberate Index

State	2020 Votes	Z-Infected	Z-Death	Z-Unemployment	Z-COVID	2020 Win Margin	Z-Risk 2-1	Z-Risk 1-1	Liberate Index
Ohio	18	-0.68	-0.37	0.80	-0.08	6.34%	-1.85	-1.32	-23.82
Minnesota	10	-1.41	-0.74	0.56	-0.53	-0.01%	-2.71	-1.64	-16.35
Georgia	16	0.26	0.00	0.96	0.41	3.53%	-0.70	-0.83	-13.25
North Carolina	15	-1.14	-0.74	-0.10	-0.66	6.75%	-1.78	-0.84	-12.58
Wisconsin	10	-0.99	-0.37	0.23	-0.37	1.17%	-1.59	-0.91	-9.13
Nevada	6	-0.40	-0.19	1.16	0.19	-4.28%	-1.74	-1.45	-8.72
New Hampshire	4	-1.43	-0.55	0.81	-0.39	0.60%	-2.79	-1.80	-7.20
Oregon	7	-1.38	-0.74	-0.32	-0.81	-7.00%	-1.79	-0.73	-5.13
Iowa	6	-0.77	-0.74	-0.40	-0.64	9.66%	-1.10	-0.35	-2.11
New Mexico	5	-0.82	-0.55	-0.41	-0.59	-5.53%	-0.96	-0.28	-1.38
Pennsylvania	20	1.36	0.73	0.97	1.02	-5.79%	1.11	0.07	1.39
Virginia	13	-0.67	-0.55	-0.99	-0.74	-1.22%	-0.23	0.38	4.97
Arizona	11	-1.08	-0.55	-1.62	-1.08	10.23%	-0.01	0.81	8.89
Colorado	9	0.22	0.18	-1.43	-0.34	-3.23%	1.84	1.63	14.70
Michigan	16	2.10	3.30	0.61	2.00	-12.10%	4.80	2.10	33.54
Florida	29	-0.37	-0.37	-1.94	-0.89	5.41%	1.20	1.57	45.59

Table 5. Z-Risk 2-1, Z-Risk 1-1, and Liberate Index

Back to the three states President Trump urged to "liberate", Michigan, Minnesota, and Virginia. In Table 5, all three states are in different groups. Michigan is in red, meaning the state is unlikely to switch sides and vote for President Trump. Minnesota is in green, indicating that the state is already on President Trump's side and will likely not switch. Virginia is in white and classified as a true swing state, though it seems to be leaning away from Trump with a positive liberate index.

Additional observations show that Florida in the red has the highest liberate index, indicating that it is the least likely out of the swing states to "swing" for President Trump, while Ohio has the lowest liberate index meaning it is most likely to vote for President Trump. Pennsylvania is a state of interest due to the large number of electoral college votes, and it is depicted as a true swing state in the white.

3.3 Visual Representation of Trump's Preferences

Figure 3 depicts the Z-COVID, Z-Unemployment, 2020 Win Margin, Z-Risk 1-1, and Liberate Index for all the swing states. President Trump prefers the Z-COVID index to be negative, indicating minimal impacts of COVID-19 on that state. Michigan is starkly positive, while Virginia and Minnesota are negative. President Trump also favors a negative Z-Unemployment, indicating a lower than average increase in unemployment is favorable. Michigan appears to be positive, so does Minnesota, while Virginia is negative. Next, President Trump would like a positive win margin, meaning his chances of winning in that state are high. Michigan is the opposite at negative, Minnesota seems to be neither positive nor negative, and Virginia appears to be barely positive. For President Trump, the Z-Risk 1-1 must be negative as well. Michigan is distinctly positive, Minnesota evidently negative, and Virginia slightly positive. Lastly, President Trump prefers a negative liberate index indicating a lower chance of the state "swinging" away from him. Michigan appears to be positive, Minnesota negative, and Virginia barely positive. Overall, Michigan seems to be dramatically the opposite of what Trump wants, while Minnesota barely leans towards President Trump, and Virginia stays virtually neutral.

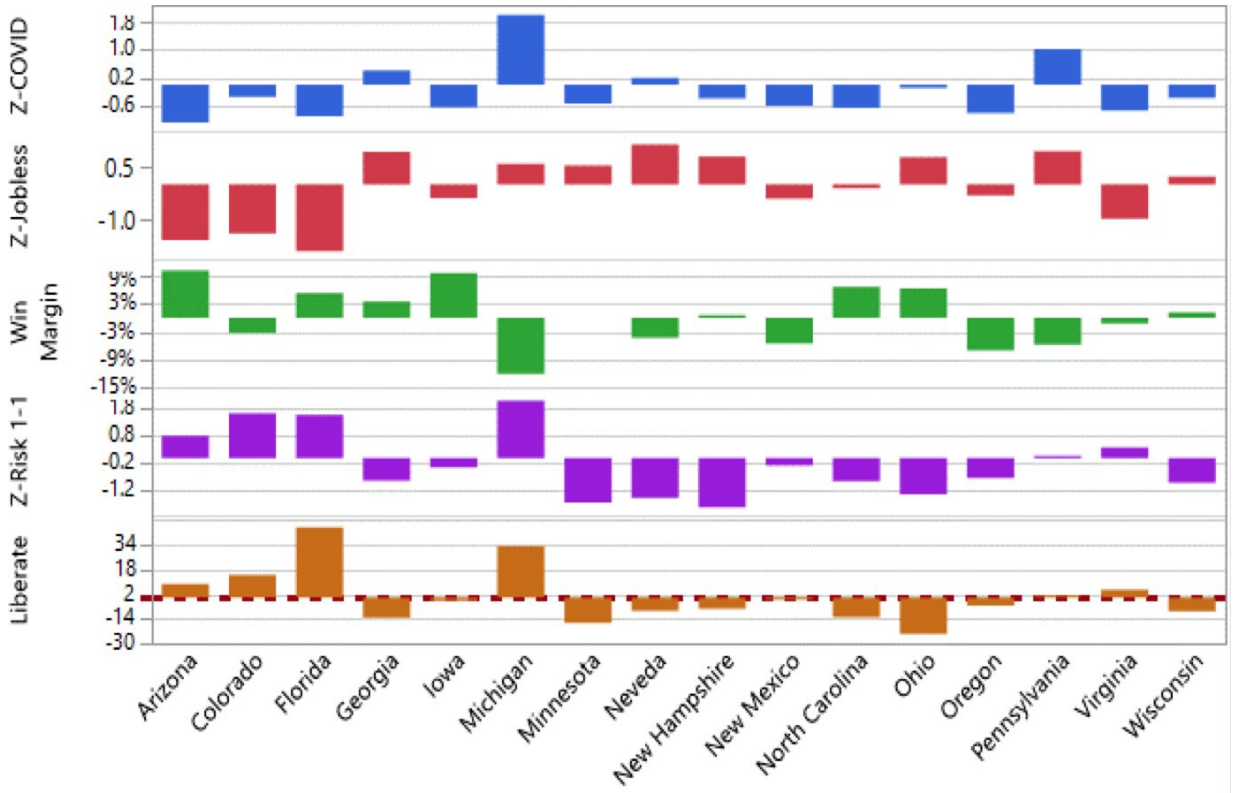

Figure 3. Multivariable Graph for Z-COVID, Z-Unemployment, Win Margin, Z-Risk 1-1, and Liberate Index

4. Conclusions

In conclusion, the authors would not suggest President Trump to liberate Michigan and Virginia because Michigan is way too risky with high infection and death rates, and if liberated the risk of spreading COVID is too high and dangerous. Virginia is hard to decide since the pros and cons are too similar and it is best left not liberated. It is not worth the risk opening just for business when unemployment rates are not a big problem in Virginia. However, President Trump could potentially liberate Minnesota, since the risk of spreading COVID is not high and there is a huge benefit for businesses and people when reducing the unemployment rate. Furthermore, President Trump could liberate Ohio, Georgia, North Carolina, Wisconsin, and New Hampshire if necessary in attempts to sway them to his side since the risk of COVID is relatively low and they could benefit from businesses reopening.

There are two possible scenarios the authors thought out. In Scenario One if President Trump doesn't push any states to liberate, he will win the 2020 election, because even though he loses 36 votes from Michigan (liberate index of 33.54) and Pennsylvania (liberate index of 1.39), he gets four more from New Hampshire (liberate index of -7.20) and

is still 7 votes above the breakeven line. He would win by 7 votes if he doesn't reopen any states. In Scenario Two, President Trump may lose Florida, which is very likely with a Liberate Index of 45.59, but win Minnesota, which is possible with a liberate index of -16.35 and barely win Virginia. In total, President Trump will still win with a single vote above the breakeven line. This may be cutting it too close for President Trump, so the authors recommend him to pay attention to Florida to maybe reconsider whether Florida should be opened in the future, however unlikely it may be. COVID-19 infection risk in Florida is quite high, so it is safer to liberate Minnesota and guarantee the votes, but it may still be risky for Virginia.

5. Future Investigations

Some future investigations could involve more scenarios and possibly a more updated version of the infected cases, deaths, and unemployment rates as these were recorded in April 2020. Additionally, the authors may monitor winning margin weekly through weekly updates on the infected cases, deaths, and unemployment rate.

Acknowledgements

We would like to thank Dr. Charles Chen for his guidance and advice on the use of statistical tools.

References

Vara, V. (2020, April 16). Coronavirus outbreak: The countries affected so far., Retrieved from https://www.pharmaceutical-technology.com/features/coronavirus-outbreak-the-countries-affected/.

W. (2020). Coronavirus Cases:. Retrieved from https://www.worldometers.info/coronavirus/?utm_campaign=homeAdUOA%3FSi.

Unemployment rate rises to record high 14.7 percent in April 2020. (2020, May 13). Retrieved from https://www.bls.gov/opub/ted/2020/unemployment-rate-rises-to-record-high-14-point-7-percent-in-april-2020.htm.

CDC. (2020, April). COVIDView Summary ending on April 18, 2020., Retrieved from https://www.cdc.gov/coronavirus/2019-ncov/covid-data/covidview/past-reports/04242020.html.

Meko, T., Lu, D., & Gamio, L. (2016, November). How Trump won the presidency with razor-thin margins in swing states. Retrieved from https://www.washingtonpost.com/graphics/politics/2016-election/swing-state-margins/.

Cillizza, C. (2012, April 16). The 9 swing states of 2012. Retrieved from https://www.washingtonpost.com/blogs/the-fix/post/the-9-swing-states-of-2012/2012/04/16/gIQABuXaLT_blog.html.

Biographies

Saloni Patel is currently a student at Stanford Online High School. She has completed JMP STIPS modules to broaden her statistical learning and usage skills. She has earned awards in local science and engineering fairs, as well as innovation challenges and business model competitions. Saloni's research interests include medicine, politics, and the environment.

Mason Chen is currently a student at Stanford OHS and serves as the student ambassador and webmaster for STEAMS. Having started STEAMS since its inception in 2014, he has held various roles such as President of the Student Chapter from 2017 to 2019. Through STEAMS, he has published more than 20 conference proceeding papers as first, second, or third author. As first author, he has won numerous awards including the Best Conference Proceeding Paper Award in the 2018 JMP Discovery Summit at age 13 as well as finishing 1st Place three times at IEOM conferences. He has also certified the IBM SPSS Statistics Level I, II, Modeler Level I, and IASSC Yellow Belt, Green Belt, and Black Belt.

Patrick Giuliano is an ASQ-certified manager of quality and organizational excellence (CMQ/OE) and holds a B.S. in Biomedical Engineering and an M.S. in Mechanical Engineering from the University of California, San Diego. He holds over a decade of experience in medical device product development, manufacturing, quality and testing - in devices ranging from the treatment of coronary/cerebral/peripheral vascular disease to heart valve dysfunction, and breast tissue reconstruction after mastectomy. He is currently a Lead Quality Engineer at Abbott Structural Heart in Menlo Park Ca, where he is the site subject-matter-expert (SME) and trainer in the application of statistics and the use of JMP software for the effective management and interpretation of engineering data using quality process methods.