

Project 1: History on Repeat? Is 2020 looking like 2016?

Levi Lee

9/22/2020

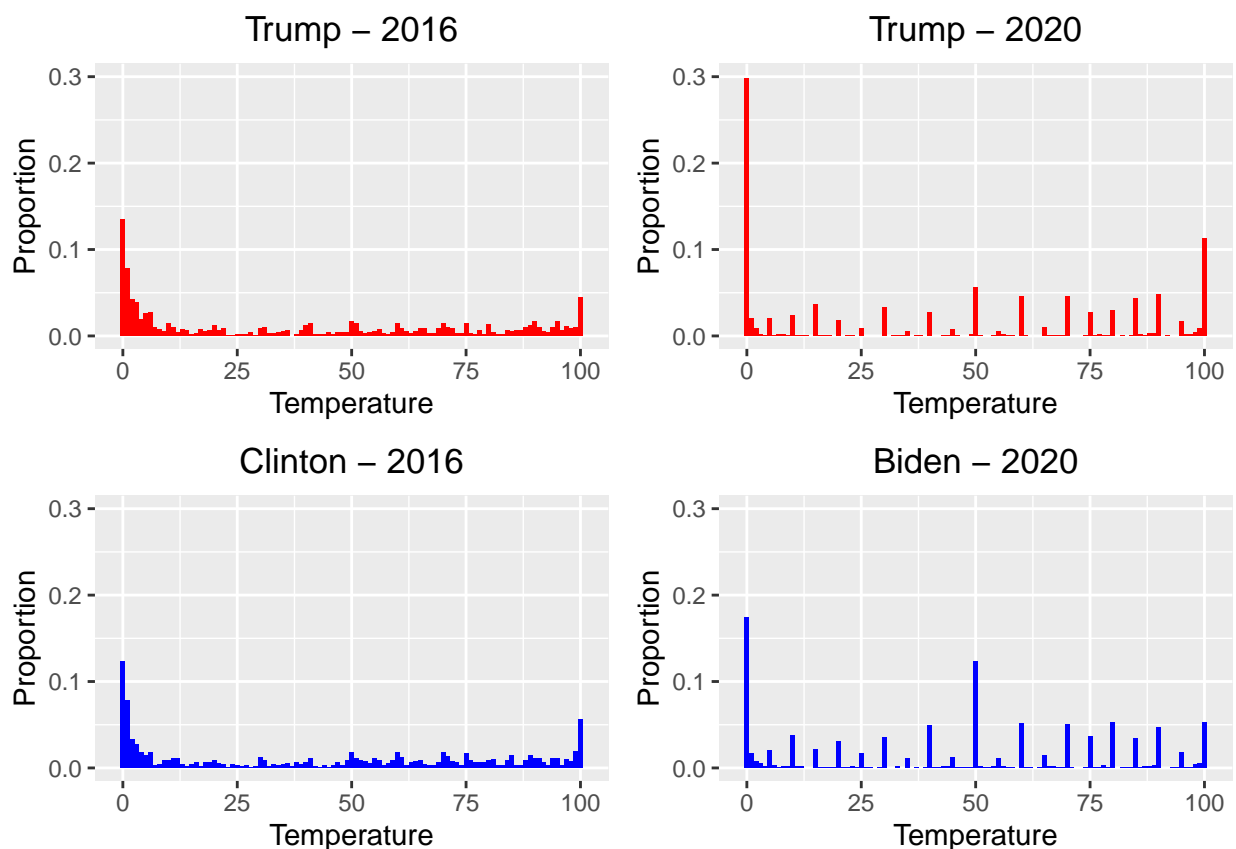
A Lesson Learned

The United States 2016 election was a perfect example of what can happen when surveys and polling data go awry. Democratic candidate Hillary Clinton was leading the polls, but only slightly so in many key states. Many of these states ended up siding with Republican candidate Donald Trump in the actual election..

A Sense of Deja-vu

We now know to take this kind of voting data with a grain of salt. In fact, as we compare the 2020 and 2016 pilot studies from American Nation Election Studies (ANES), we see a similar story.

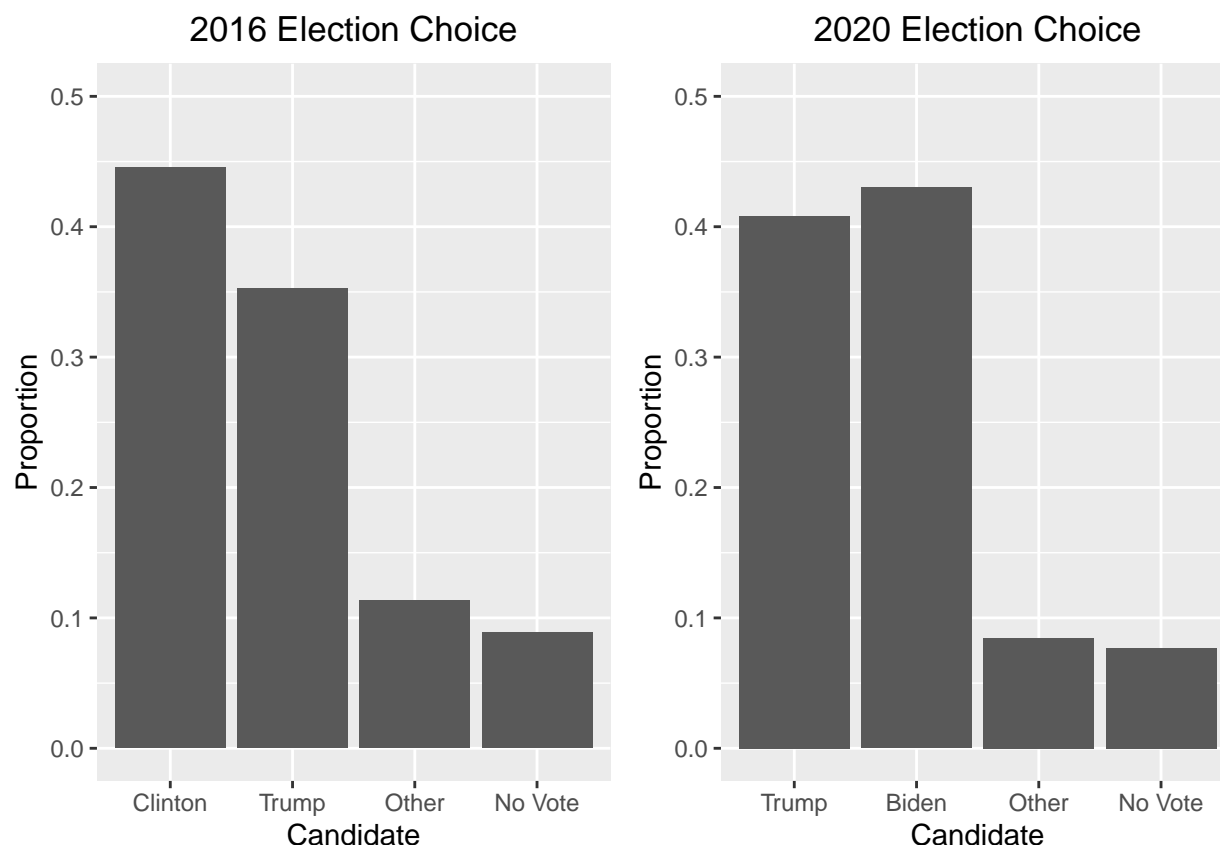
Using the notion of feeling thermometers developed by ANES, the barplots above reflect the full range of respondents' feelings toward the Republican and Democratic candidates of 2016 and 2020—Donald Trump (R-2016, R-2020), Hillary Clinton (D-2016), and Joe Biden (D-2020). These feelings are in the form of a temperature ranging from 0 to 100, where 0 degrees indicates a very cold or unfavorable feeling, 100 degrees indicates a warm or favorable feeling, and 50 degrees indicates neither favorable nor unfavorable.



In 2016, we see mainly opposition to both candidates, but Clinton was ahead with an average of 42.99 degrees compared to Trump’s average of 38.38 degrees.

What about 2020? Trump’s 2020 barplot, while more polarized, looks similar to the one in 2016 in terms of overall shape. Meanwhile, Biden’s plot, while a lot more neutral compared to Clinton in 2016, still has both major opposition among its respondents. The numbers are similar as well—Biden is leading with a 45.17 degree average compared to Trump’s 42.42 degree average.

The similarities seem to continue even further. From the barplots below, we see that, back on 2016, when asked *“If the 2016 presidential election were between Hillary Clinton for the Democrats and Donald Trump for the Republicans, would you vote for Hillary Clinton, Donald Trump, someone else, or probably not vote?”*, the responses showed that Clinton came up on top over Trump by a noticeable margin. Now, in 2020, when asked *“If the 2020 presidential election were between Donald Trump for the Republicans and Joe Biden for the Democrats, would you vote for Donald Trump, Joe Biden, someone else, or probably not vote?”*, the responses show that Biden comes up on top as well.

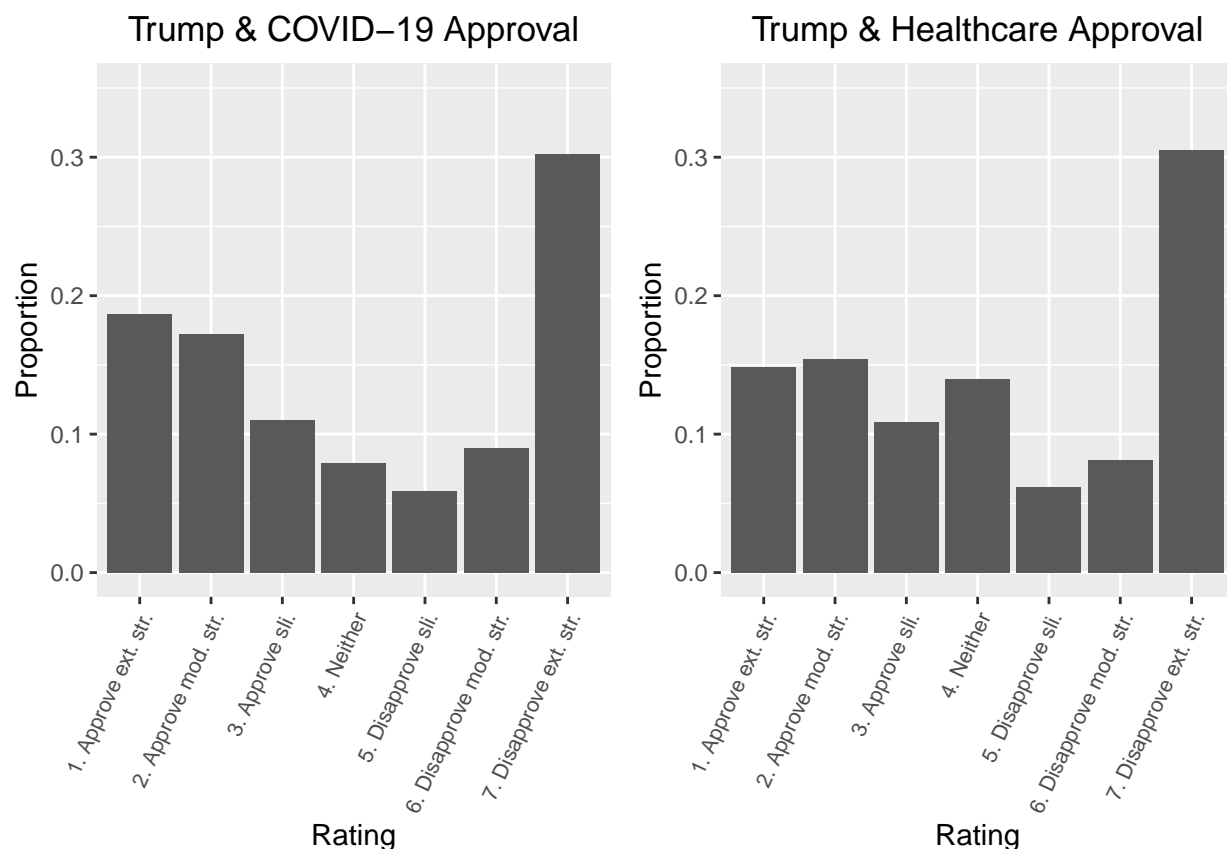


It is tempting to use 2016 as justification for doubting what we see in this data, and while the results of 2016 are important to keep in mind, we should be careful in jumping to conclusions. Doing so would only cause us to make the same mistake as we did in 2016. By considering current events and what is different about 2020—the good and the bad—we see that they have a significant effect on public opinion and therefore how we view this upcoming election as well. We present three different justifications below using additional public opinion data from ANES that was contained in the 2016 and 2020 pilot data studies.

COVID-19

Perhaps the most obvious difference between 2016 and 2020 is the issue of COVID-19 in the United States. While the the pandemic has not only shutdown many business and the economy, at the time of this writing, it has claimed over 200,000 lives in this country alone. And throughout the pandemic, arguments about Trump’s involvement, the use of masks, symptoms, and how it spreads continue to rage on.

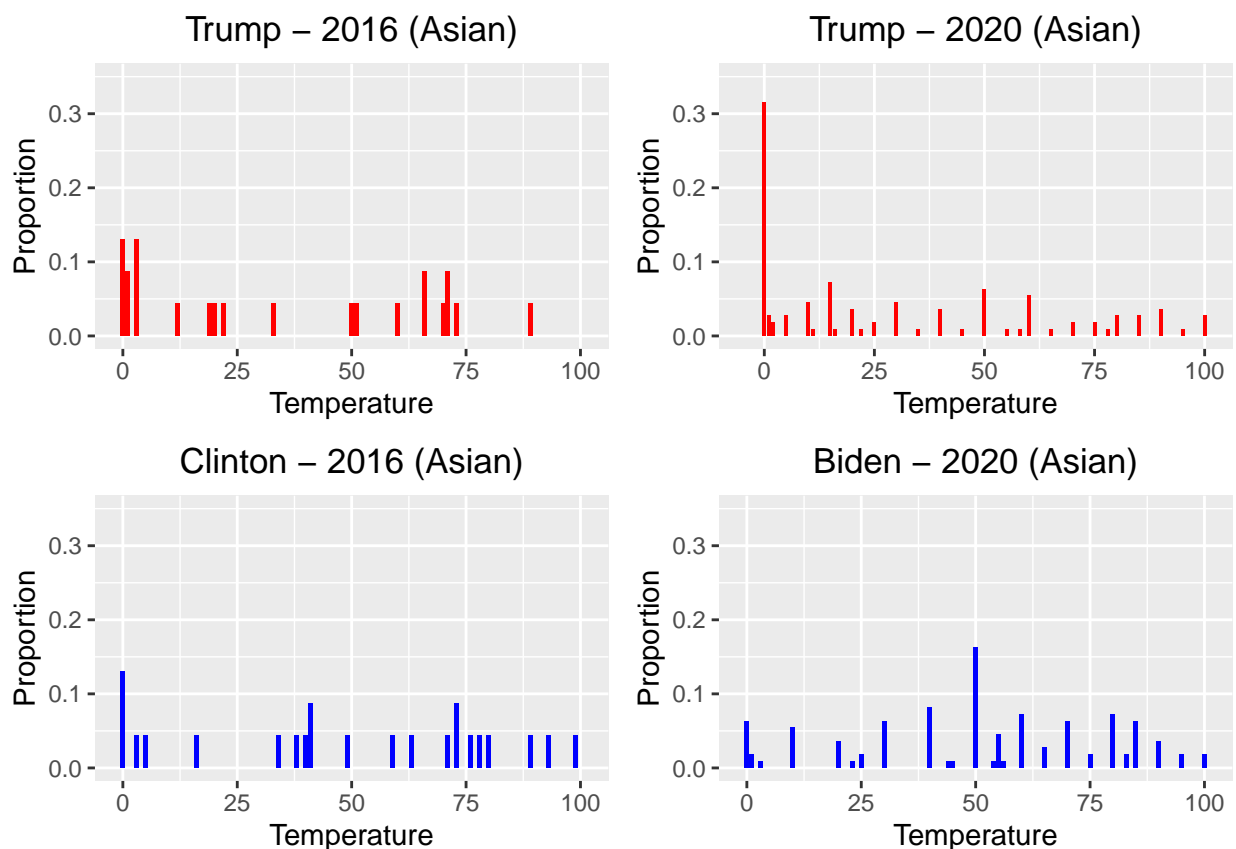
Responses to questions unique to ANES’s 2020 Pilot Survey—*“Do you approve, disapprove, or neither approve nor disapprove of the way Donald Trump is handling the coronavirus (COVID-19) outbreak?”* and *“Do you approve, disapprove, or neither approve nor disapprove of the way Donald Trump is handling health care?”*—are shown below.



A majority seem to believe that Trump has handled both issues rather poorly, which is not only a major factor in the political scene of 2020, but it potentially hurts his chances for re-election.

Asian Americans in the Public Eye

Additionally, due to COVID-19 originating from China, there have been major tensions among Asian-Americans and Trump; to make matters worse, many Asians and Asian-Americans have been harassed or discriminated against as the virus spread. The president's insistence in calling the disease "the Wuhan virus" affects not only China's perspective of Trump but the United State's voting population of Trump as well—particularly Asian-Americans.

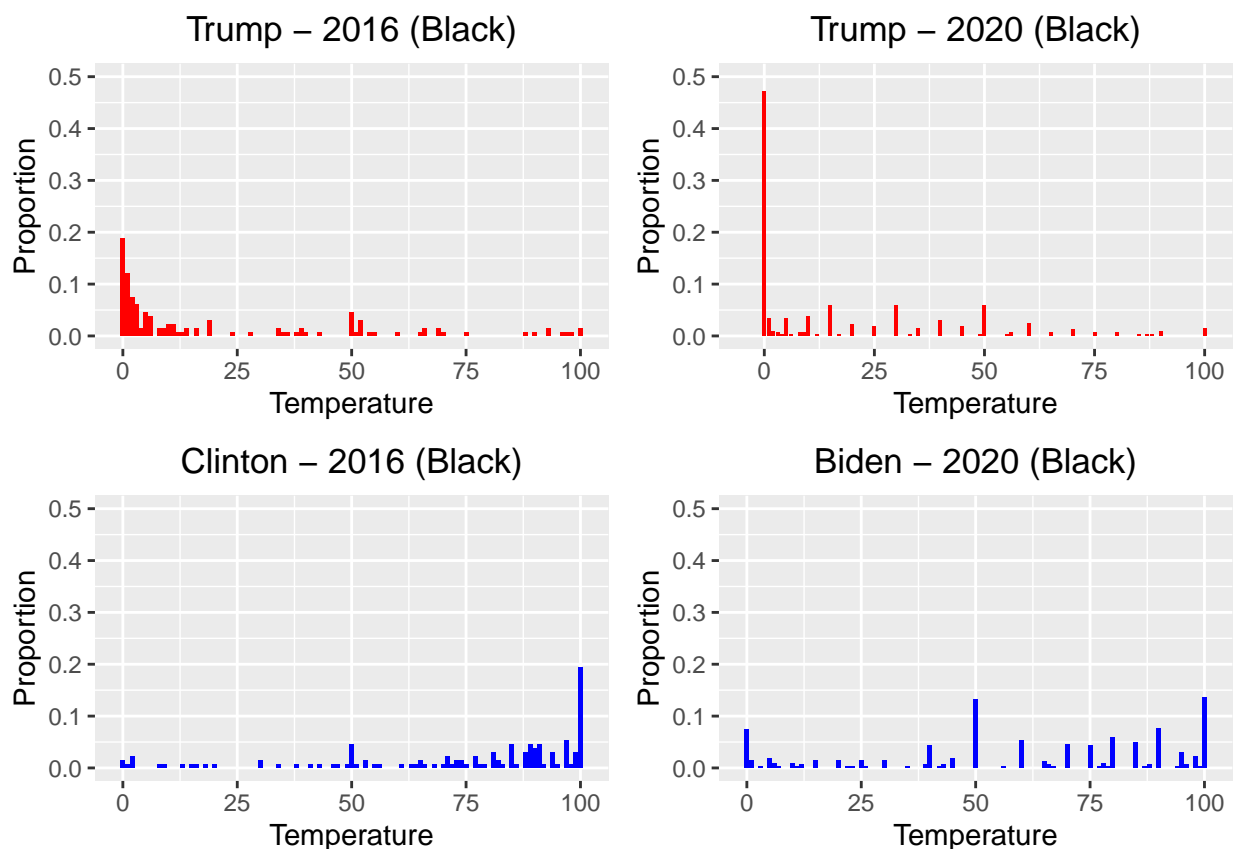


Let's break down the four barplots from above and focus only on the Asian-American voters. We see opinions on Trump have been more defined much more unfavorable in 2020 compared to 2016. The average has decreased from 34.09 degrees in 2016 to 29.43 degrees in 2020. In 2016, about 13% of Asian-American voters sampled were extremely disapproving with Trump; that number has risen to over 30% in 2020.

Opinions on Biden are much more middling in 2020 compared to Clinton in 2016 but both positive overall. With Biden's average of 50.7 degrees, he is a few points ahead compared to Clinton's 48.74 degree average back in 2016.

Black Lives Matter

In 2020, we have seen the deaths of many Black lives including George Floyd and Breonna Taylor, and opinions on police brutality and racism continue to clash among political parties and other groups. However, Black Lives Matter is not a recent movement, dating as far back as 2013, and many injustices against Black lives in the United States have been committed prior to this year. Because of this, opinions on Trump—as well as his response to the movement—continue to change.



Now let's break down the four barplots we introduced again, but this time, we focus on only Black voters. Within this particular group, we again see significant changes between Black voters' opinions on Trump. Overall, feelings toward Trump have become much colder, with over 40% of surveyed Black respondents saying that they are completely unfavorable with Trump compared to just under 20% in 2016.

Overall, the average has decreased from 21.35 to 17.45 degrees. It is also worth noting that, at the time of this data, the average for Biden is 74.49 while Clinton in 2016 was 60.64, which indicate moderate favorable feeling toward each candidate for their respective years.

A Sense of Optimism

Regardless of which candidate people want to vote for, the optimism mentioned in the subheading is referring to one's ability to trust in the data. Dismissing this data because of the outcome of 2016 is not only rash and misleading but very dangerous. It takes away the very power that Americans have the right to exercise to make changes in ways that matter to them.

However, data should never be analyzed without context from major current events, especially if data—such as ANES's—records public opinions on such events. This article is intended to give readers of just a few examples of how 2020 is unlike 2016, but many more can be found. However, the data is still subject to many types of biases, which we outline a few below. And even with the most updated data, only time will tell who the winner of the election will be.

A Disclaimer - Notions of Bias

Polling and election data and surveys remain an important part of the election process and serves as a way to inform the public. However, like most non-experimental data, it can be difficult to draw concrete conclusions

and generalize to a population due to many different types of biases. We note some below.

Historical bias: Data immediately becomes old the moment we record it; information continues to change and while there is no way to overcome this without getting newer data, it is important to be mindful of this data in the context of not only current events, but *recent* events. The conclusions we drew are only relevant if current suggests that it is.

Sampling and representation biases: ANES surveys may not necessarily reflect or generalize to the population of interest. This is especially the case with minority groups as they tend to be undersampled. One way to deal with this (as did ANES) is to intentionally oversample a particular minority group, but must weigh the responses appropriately in order to generalize to a population. While we try to avoid sweeping generalizations whenever possible, we still caution readers and other analysts to treat the data with a sense of caution.

Convenience, voluntary and non-response biases: The study was mainly done online, which could affect the type of people who actually get recorded among the responses. There may be a percentage of the population that cannot or do not know how to use a computer, for example. Some many choose not to do it or leave certain questions blank. Respondents were also paid in some way for their responses, which could have an influence on the subset that is being sampled as well. This is not something that our analysis accounted for, but again, we remind readers and analysts to proceed with caution.

Measurement error and missing data: Some data entries may be incorrect or even missing entirely, and it is up to the data scientist to decide how it should be handled. They could choose to omit that particular row from the analysis or try to seek out the correct information if they have the means to do so. There is no way for us to verify the accuracy of this data, we would like to acknowledge that missing data was omitted for our analysis.