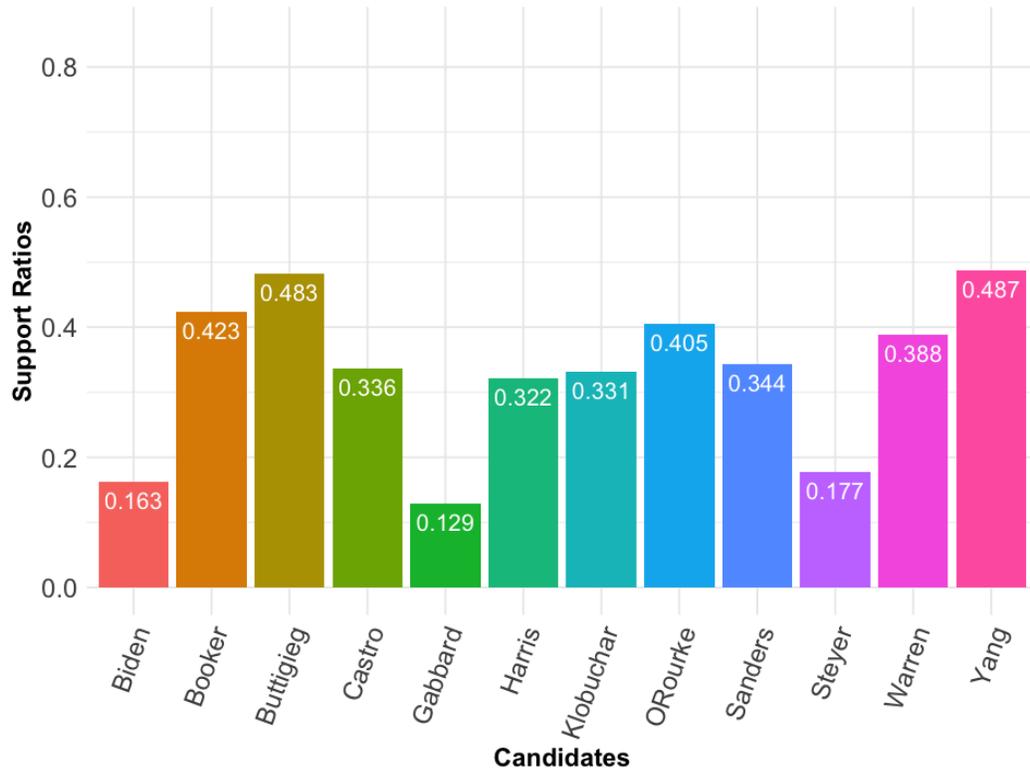


Yang and Buttigieg the Big Winners from Last Night's Debate According to Twitter

The Center for Survey Research (CSR) at East Carolina University collected more than 240,000 tweets (of which 129,019 could be categorized as positive or negative for a specific candidate) made during the hours of the Democratic national debate on October 15, 2019. Using sentiment analysis to calculate a ratio of positive to negative tweets for each of the ten candidates, Andrew Yang and Pete Buttigieg came out on top. (For additional information about the analysis, see the methodology section below.)

| Candidate | Positive support ratio | Total categorized tweets |
|------------------|------------------------|--------------------------|
| Andrew Yang | .487 | 13,068 |
| Pete Buttigieg | .483 | 12,092 |
| Cory Booker | .423 | 7,691 |
| Beto O'Rourke | .405 | 6,041 |
| Elizabeth Warren | .388 | 24,032 |
| Bernie Sanders | .344 | 14,051 |
| Julian Castro | .336 | 4,019 |
| Amy Klobuchar | .331 | 7,309 |
| Kamala Harris | .322 | 10,697 |
| Tom Steyer | .177 | 4,056 |
| Joe Biden | .163 | 15,921 |
| Tulsi Gabbard | .129 | 10,041 |



As shown above, the results that emerge from the sentiment analysis of tweets during last night’s debate reveal distinct winners and losers. First, Andrew Yang and Pete Buttigieg received the most positive sentiment from Twitter during Tuesday night’s debate. Both Yang and Buttigieg also generated a reasonable degree of interest, with a total number of 13,068 tweets and 12,092 tweets respectively. Only the front runners, Elizabeth Warren (24,032), Joe Biden (15,921), and Bernie Sanders (14,051) had a higher number of tweets (although the sentiment of those tweets was less positive when compared to Yang and Buttigieg). Cory Booker generated the third most positive ratio of tweets during the debate, but had only 7,691 tweets, which ranked eighth out the twelve candidates.

If there was an identifiable loser on Twitter, it was Tulsi Gabbard and Joe Biden, both of whom generated the least positive reaction to their respective debate performances. As one of the top candidates in recent polls, it is not surprising to see that Biden had the second highest number of tweets of any candidate. However, the tweets about Biden were disproportionately negative, with only Tulsi Gabbard scoring lower in the sentiment analysis.

Our results parallel some the post-debate analysis from political writers and pundits who likewise identified [Yang and Buttigieg as winners in Tuesday night’s debate](#), but not all. Others placed [Warren and Sanders](#) in the top two. Our analysis of Twitter users placed Warren and Sanders closer to the middle of the pack.

Although Twitter users are certainly not representative of the nation’s population as a whole, it does reflect an audience of people who are engaged enough politically to voice their opinions. In

primary elections, where voter turnout is often low, this engaged audience seems worthy of some attention.

If the larger winds of public opinion do happen to blow in Twitter's direction, then Yang and Buttigieg will be the beneficiaries. In contrast, Gabbard's presidential campaign did little to help itself among Twitter users on Tuesday night, while Joe Biden struggled as well to impress the Twitter-verse.

* * * * *

About the Authors

Dr. Peter L. Francia is the Director of the Center for Survey Research and is a professor of political science at East Carolina University. He can be reached at 252-328-6126 or at franciap@ecu.edu.

Dr. Baekkwon Park is a Senior Data Analyst at the Center for Survey Research at East Carolina University.

Dr. Venkat N. Gudivada is chairperson and professor in the Department of Computer Science at East Carolina University.

Jennifer Andriot is a graduate student in data science in the Department of Computer Science at East Carolina University.

Methodology Summary

The analysis in this article is based on a collection of more than 240,000 tweets containing hashtags #demdebate, #demdebates, #democraticdebate, #democraticdebates, #presidentialdebate, and #presidentialdebates from Twitter between 8:00 – 11:00 p.m. on October 15, 2019, through twitter streaming API. To generate our results, we filtered the data so that the analysis contained only tweet texts for the each of the twelve Democratic primary candidates. The analysis relied on a supervised machine learning approach (Support Vector Machine, accuracy 83.1). This required annotated training data. We manually annotated about 4,000 tweets. Five groups of students were assigned a random set of tweets and manually labeled each tweet into two classes: positive or negative. In order to check the inter-coder reliability, we examined all the labeled data again. After retrieving and filtering all the tweets about the primary debate, we pre-processed the texts to eliminate noisy data. After classifying the tweets into two categories (positive or negative), we calculate the proportion of positive tweets over the total tweets for each candidate.