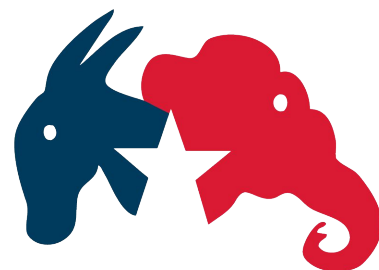# PEORIA Project
2020 Election Predictions Model Results

New Hampshire Primary - Democratic Party

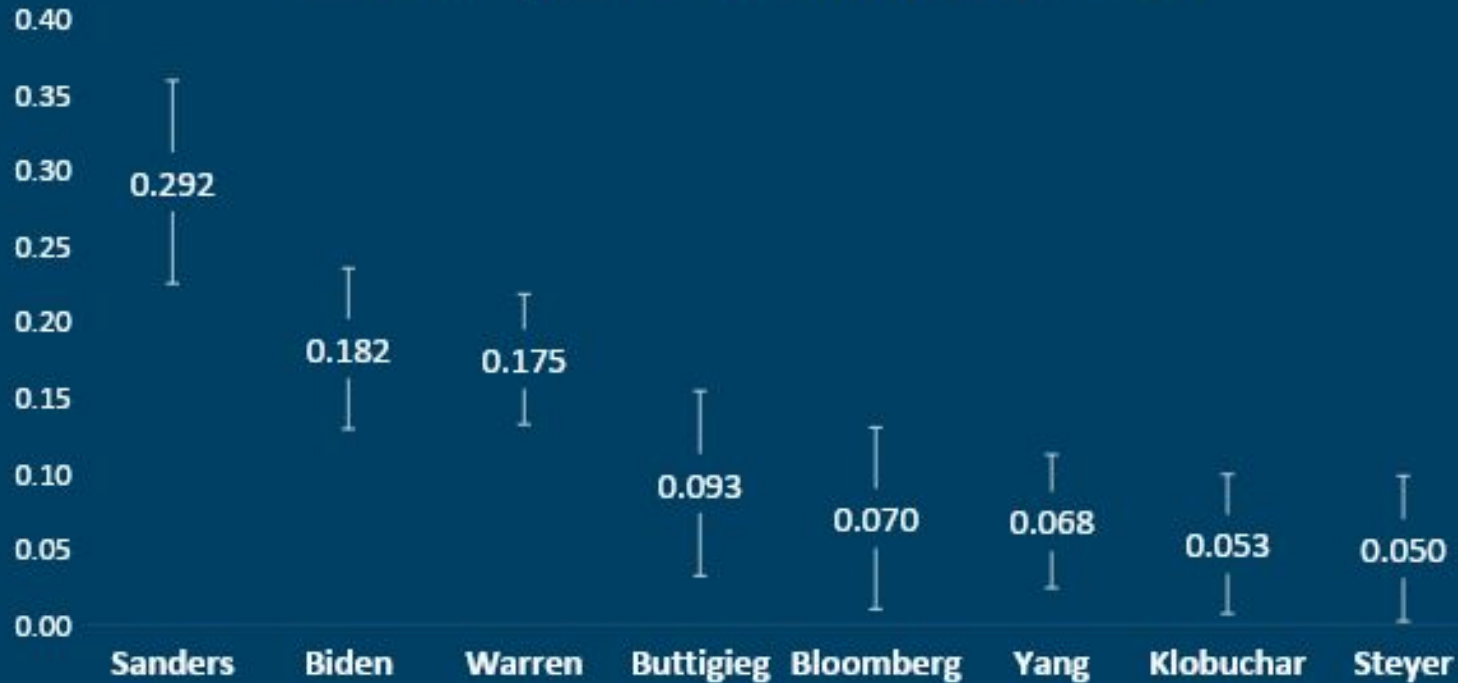Sanders, Biden, and Warren Likeliest to Win, In That Order

Innovative Model Incorporates Social Media Variable
of Twitter Mentions to Yield Ranges of Likely Results

Results

New Hampshire Vote Share Prediction

Candidate	Average Predicted Vote Share	Lower Bound	Upper Bound
Sanders	**0.292**	0.226	0.359
Biden	**0.182**	0.129	0.235
Warren	**0.175**	0.131	0.218
Buttigieg	**0.093**	0.032	0.154
Bloomberg	**0.070**	0.009	0.130
Yang	**0.068**	0.024	0.111
Klobuchar	**0.053**	0.008	0.099
Steyer	**0.050**	0.001	0.098

The chart and table report the predicted primary vote share in New Hampshire for each candidate. For example, Bernie Sanders is predicted to receive 29.2% of the vote share. The bars indicate the upper and lower bounds for the prediction (95% confidence interval).
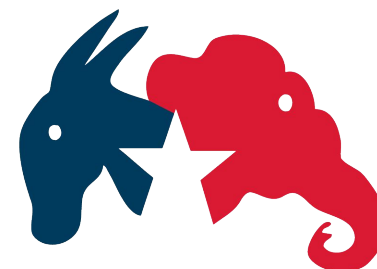
Our Key Three Variables

Our model predicts a candidate's performance based entirely on three factors: Twitter mentions, cash on hand, and endorsements.

While we are aware that in important ways the Twitter universe does not necessarily reflect the electorate, the quantity of **Twitter Mentions** is a good proxy for the "buzz" a candidate is getting within the wider electorate, and reflects the activity of important opinion leaders.

Cash on Hand reflects the strength of the candidate in the "money primary."

Endorsements indicate each candidate's strength within the party, which speaks to the debate over whether the party decides the outcome of the nomination. Endorsements are not statistically significant but are an important control variable in our model.
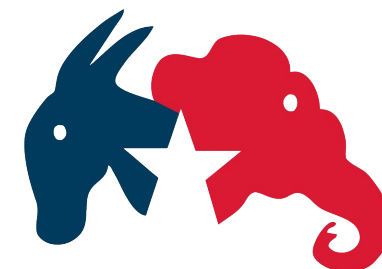
Explanation of Model

What Our Model Does

Our model predict the Iowa caucus vote share for each Democratic candidate using three predictor variables in an Ordinary Least-Squares (OLS) multiple regression.

How We Predict Vote Share

In order to predict each candidate's vote share, we input the latest variable data (see below) into the regression model to generate an estimate as well as an upper- and lower-bound.

Cash on Hand: Measured as a percentage share of the total cash on hand for all candidates within the party. The most recent data were for Quarter 4 of 2019. Source for data: FEC.gov

Twitter Mentions: Measured as the number of mentions on Twitter for each candidate as a percentage share of the total number of mentions for all candidates within the party. The data for these models were from the month of January, 2020 and represent only tweets generated in the United States and in English. Because Biden was often mentioned in conjunction with Donald Trump's impeachment, rather than his candidacy, the following words were excluded from tallying his total mentions: "Hunter", "son", "Burisma", "impeachment", "corruption" and "Ukraine". Source for data: Crimson Hexagon

Endorsements: Measured as the total number of endorsements for each candidate by US Senators, members of the US House of Representatives, former Presidents and Vice Presidents, former presidential candidates from the current election cycle who had dropped out of the race, and elected statewide officials, state legislative leaders, and mayors of large cities. The data for these models were updated the month of February, 2020. Source for data: FiveThirtyEight.com

How We Chose Our Model

To find the best fitting model, we used campaign data from 2012 and 2016. Several models were created, including OLS regression, longitudinal regression (using Q1 through Q4 cash on hand as well as monthly twitter mentions), ridge regression, partial least squares regression, and principal component regression. Regressions were run with and without incumbent candidates' parties considered. The model with the lowest RMSE while maintaining the highest possible R^2 was chosen for
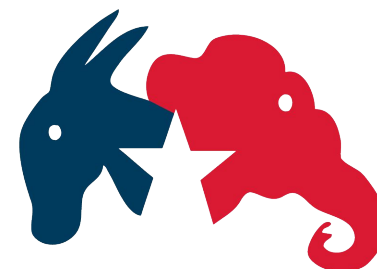
Descriptive Table of Variables and Regression Model for New Hampshire Vote Share

Means, Standard Deviations, and Correlations

Variable	M	SD	1	2
1. Twitter Mentions	0.148	0.235		
2. Cash on Hand	0.146	0.246	0.846	
3. Endorsements	137.556	257.938	0.293	0.449

Summary of Regression of the New Hampshire Primary Vote Share Prediction

Variable	Estimate	SE B	β
Intercept	0.02121	0.01671	
Twitter Mentions	0.57450**	0.11073	0.66489
Cash on Hand	0.27813*	0.11322	0.33688
Endorsements	-0.00005	0.00006	-0.06088

adj R^2 = 0.8823, $F(3,23)$ = 65.97

$*p < 0.05.$ $**p < 0.01.$

This table demonstrates that Twitter Mentions and Cash on Hand are positive significant predictors of a candidate's vote share in the New Hampshire Primary.
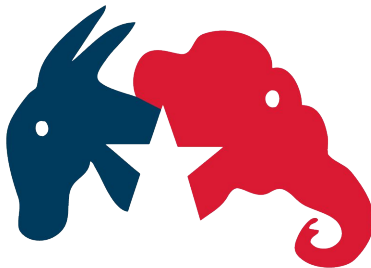
Thanks for reading!
Come back each week for new predictions!

Questions may be directed to the authors:

Meg O'Neill
megoneill@gwu.edu
Data Scientist for The PEORIA Project,
Graduate School of Political Management,
The George Washington University

Michael Cornfield
corn@gwu.edu
Associate Professor and Director of The PEORIA Project,
Graduate School of Political Management
The George Washington University

Todd Belt
tbelt@gwu.edu
Professor and Director of the
Political Management Master's Program,
Graduate School of Political Management
The George Washington University

Lara Brown
larambrown@gwu.edu
Associate Professor and Director of the
Graduate School of Political Management,
The George Washington University

Appendix: Twitter Data

2012 Candidate	January Mentions	Share within Party
Ron Paul	420,396	0.304
Mitt Romney	360,698	0.261
Newt Gingrich	250,634	0.181
Rick Santorum	169,448	0.122
Rick Perry	115,263	0.083
Barack Obama	92,845	0.992
Jon Huntsman	36,999	0.027
Michele Bachmann	30,058	0.022
Randall Terry	603	0.006
Fred Karger	570	0.000
Darcy Richardson	99	0.001
John Wolfe Jr	69	0.001
Bob Ely	16	0.000

2016 Candidate	January Mentions	Share within Party
Donald Trump	1,429,143	0.524
Bernie Sanders	726,110	0.499
Ted Cruz	710,975	0.261
Hillary Clinton	697,982	0.480
Marco Rubio	142175	0.052
Ben Carson	107,806	0.040
Jeb Bush	104,689	0.038
Rand Paul	84,167	0.031
Chris Christie	48,247	0.018
Carly Fiorina	34,643	0.013
Martin O'Malley	30,336	0.021
John Kasich	28422	0.010
Rick Santorum	17,384	0.006
Mike Huckabee	16,125	0.006
Jim Gilmore	4,132	0.002

2020 Candidate	January Mentions	Share within Party
Donald Trump	3,961,690	0.997
Bernie Sanders	2,083,893	0.331
Joe Biden	1,995,614	0.317
Elizabeth Warren	1,079,934	0.171
Andrew Yang	337,386	0.054
Pete Buttigieg	248,416	0.039
Tulsi Gabbard	166,960	0.026
Amy Klobuchar	160,707	0.025
Tom Steyer	107,706	0.017
Michael Bloomberg	102,035	0.016
Michael Bennet	10,218	0.002
Deval Patrick	9,666	0.002
Joe Walsh	9,163	0.002
Bill Weld	3,982	0.001
Rocky De La Fuente	100	0.000

Republican candidates in red, Democratic candidates in blue.
For more information: The PEORIA Project, and the 2020 Weekly Tweeterboard.
For a discussion of the differences between the social media electorate and the broader electorate, click here.