DATA-DRIVEN RED TEAMING

FINAL RESULTS

Group 6

Drew Adams, Anuska Mondal, Eric Rash, Sai Somana

Data & Variables

We collected our data from the following sources:

- MIT Election Lab
- Harvard Database
- American Presidency Project
- St. Louis Federal Reserve
- World Bank
- National Geographical Information System

Then, we selected our variables based on what we determined to be most relevant and useful for our regression model. To the right are our selected variables used in the final regression.

- Democratic Senate vote share
- Open seat
- Incumbent party
- Incumbent democrat
- Unemployment change
- Democratic presidential vote share
- Party president
- President republican
- Approval
- Inflation
- Proportion of urban

Building the Regression Model

Democratic vote share = β_0 + β_1 Democratic presidential vote share + β_2 Urban proportion + β_3 (Unemployment change x Incumbent Democrat) + β_4 Incumbent Democrat + β_5 (Inflation x President Republican) + β_6 (President Republican x Approval) + β_7 President Republican

Coefficients				
	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-0.122	0.067	-1.776	0.0761
dem_pres_vote_sharre	0.444	0.056	7.968	6.17e-15***
propurb	0.017	0.032	0.537	0.591
unemp_change	-0.006	0.006	-0.951	0.342
inc_d	0.168	0.011	15.338	<2e-16***
inflation	0.286	0.002	4.789	2.03e-06***
pres_r	0.286	0.069	4.171	3.40e-05***
approval	0.005	0.001	4.054	5.58e-05***
unemp_change : inc_d	-0.010	0.008	-1.308	0.191
inflation : pres_r	0.011	0.005	2.313	0.021*
pres_r : approval	-0.005	0.001	-4.105	4.50e-05***

Initial Results

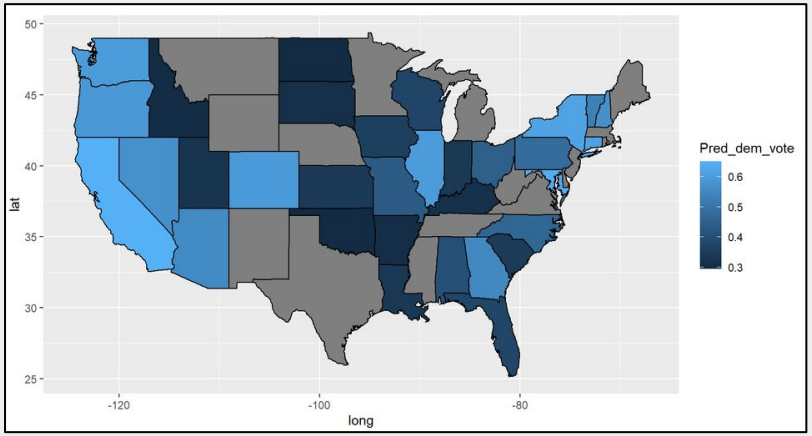
Democratic Wins:

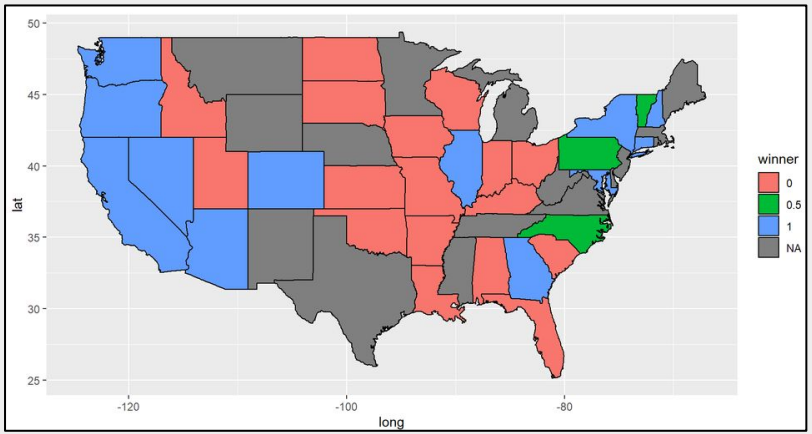
- Arizona
- California
- Colorado
- Connecticut
- Georgia
- Hawaii
- Illinois
- Maryland
- New Hampshire
- Nevada
- New York
- Oregon
- Washington

Republican Wins:

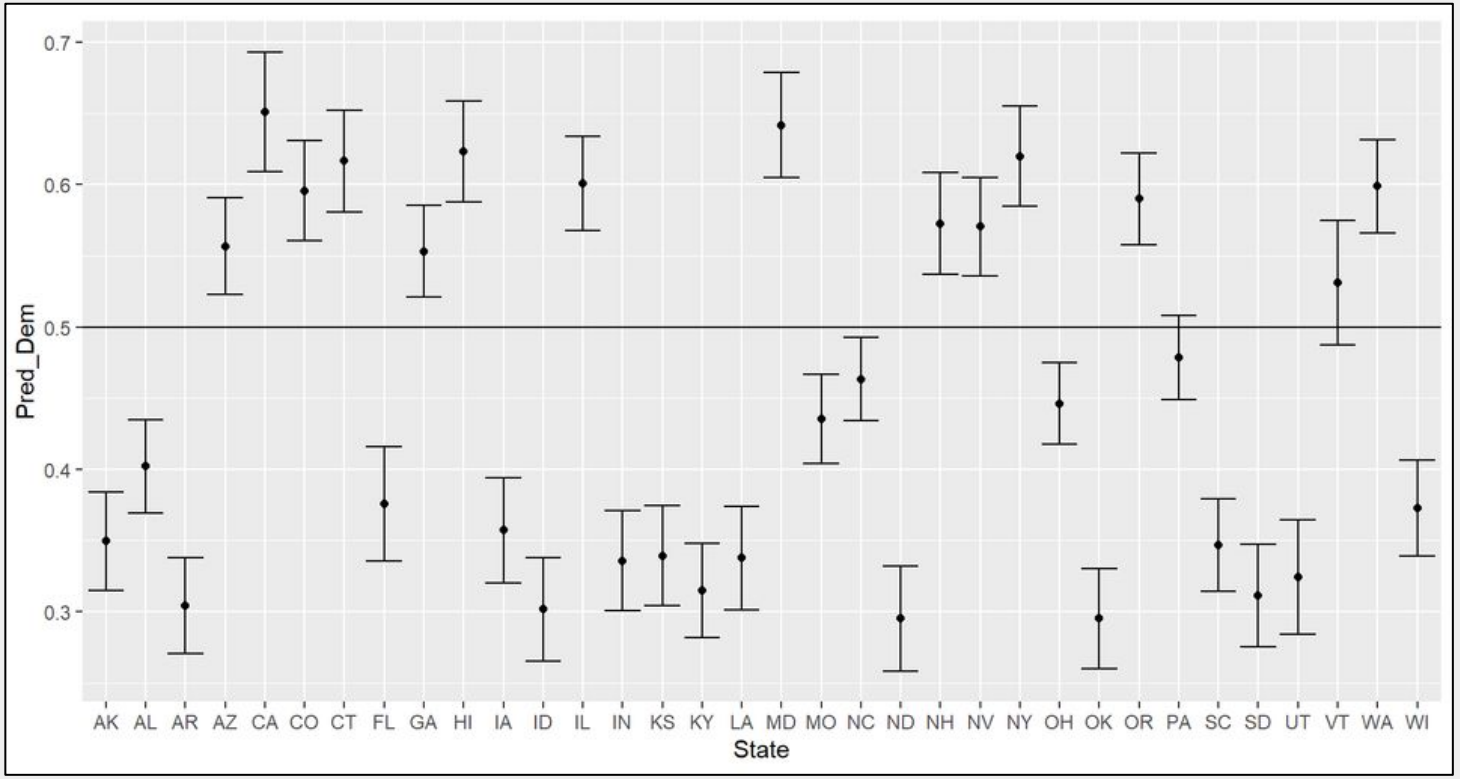
- Alaska
- Alabama
- Arkansas
- Florida
- Iowa
- Idaho
- Indiana
- Kansas
- Kentucky
- Louisiana
- Montana
- North Carolina
- North Dakota
- Oklahoma
- South Carolina
- South Dakota
- Utah
- Wisconsin

Toss-Ups?

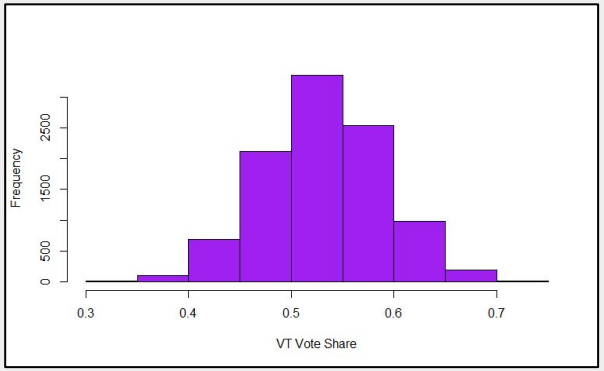


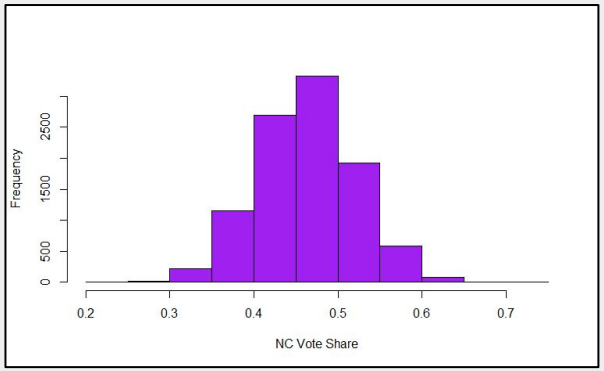


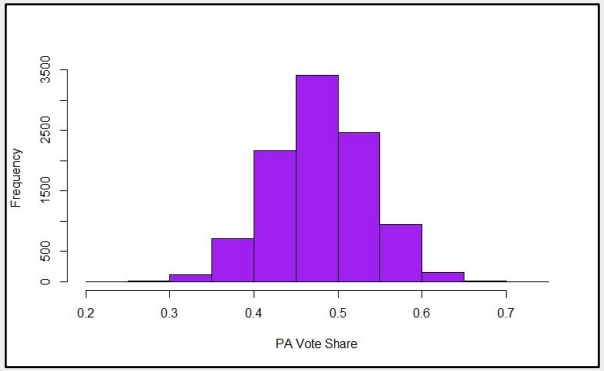
Confidence Intervals



Simulating Toss Ups





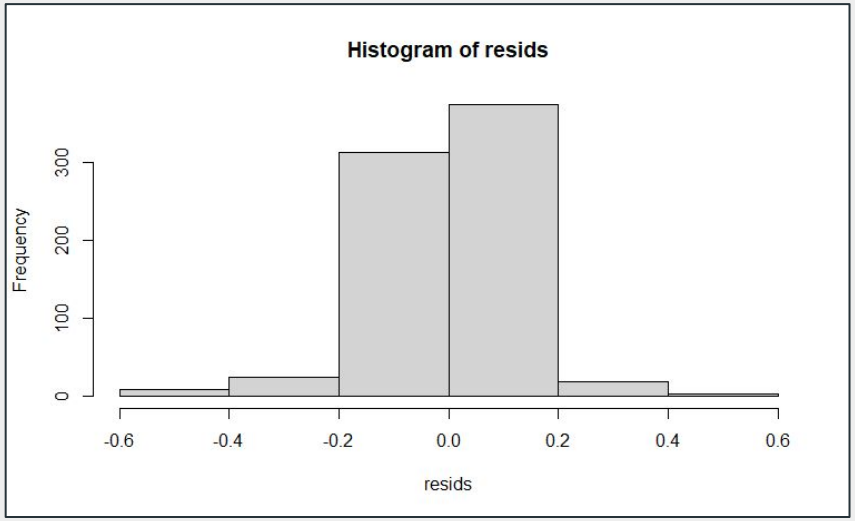


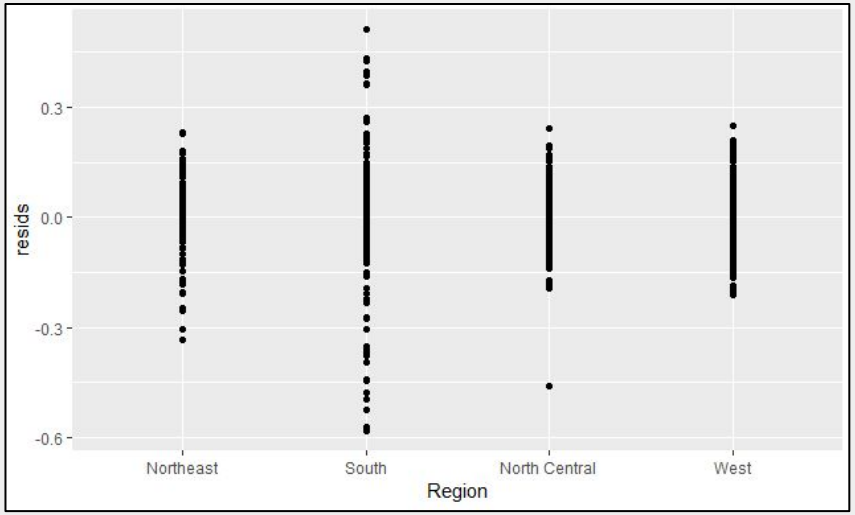
Vermont:
Democrat

North Carolina:
Republican

Pennsylvania:
Republican

Residuals

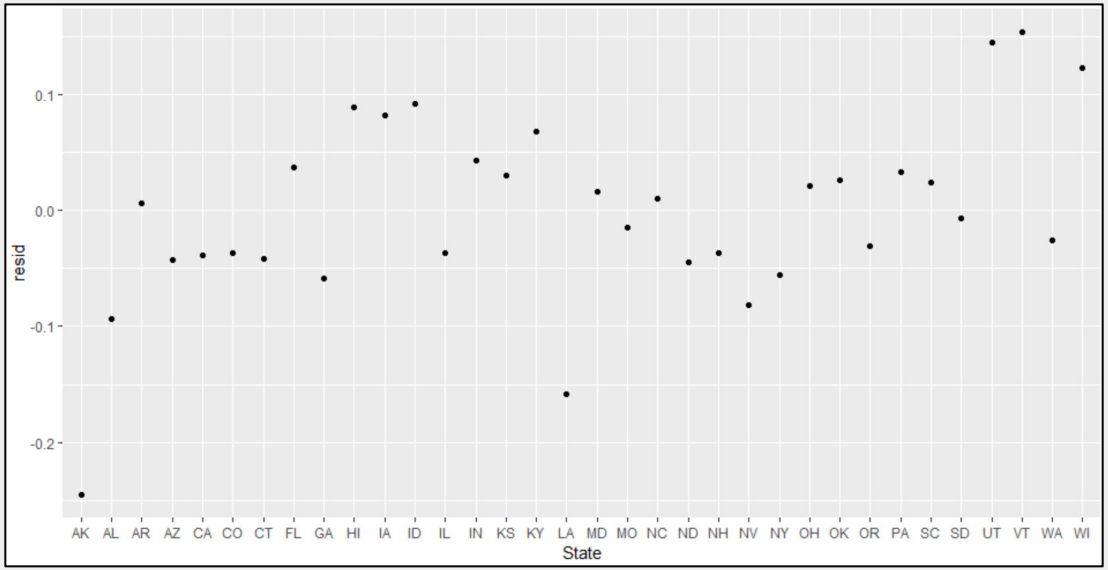




The plots above show a histogram of our residuals and these residuals disaggregated by region in the U.S.

Comparing Results

While our basic predictions (which party would win) were mostly true, our precise predictions (the actual percentage of the vote Democrats would receive) were not so in several cases. The residuals derived from subtracting our predicted results from the real results are plotted below:



State	Predicted Outcome	Actual Outcome	State	Predicted Outcome	Actual Outcome
Alaska	Rep.	Rep.	Maryland	Dem.	Dem.
Alabama	Rep.	Rep.	Missouri	Rep.	Rep.
Arkansas	Rep.	Rep.	North Carolina	Toss-Up[2]	Rep.
Arizona	Dem.	Dem.	North Dakota	Rep.	Rep.
California	Dem.	Dem.	New Hampshire	Dem.	Dem.
Colorado	Dem.	Dem.	Nevada	Dem.	Dem.
Connecticut	Dem.	Dem.	New York	Dem.	Dem.
Florida	Rep.	Rep.	Ohio	Rep.	Rep.
Georgia	Dem.	Run-off	Oklahoma	Rep.	Rep.
Hawaii	Dem.	Dem.	Oregon	Dem.	Dem.
Iowa	Rep.	Rep.	Pennsylvania	Toss Up	Dem.
Idaho	Rep.	Rep.	South Carolina	Rep.	Rep.
Illinois	Dem.	Dem.	South Dakota	Rep.	Rep.
Indiana	Rep.	Rep.	Utah	Rep.	Rep.
Kansas	Rep.	Rep.	Vermont	Toss Up	Dem.
Kentucky	Rep.	Rep.	Washington	Dem.	Dem.
Louisiana	Rep.	Rep.	Wisconsin	Rep.	Rep.

Conclusion

Model: Highly Effective

Of course, there are discrepancies, so what could we have done better?

- ★ Variable selection
  - Demographic information for candidates. I.e. Gender, race, education, etc.
  - Campaign Funding
  - Political Endorsements