

Presidential Election



Predicting 2016 Presidential
Election Outcome by County

by Miranda Zhao

Background

"...What happened in the Election data and algorithms.. is a combination of

the shortcomings of
polling, analysis and
interpretation

- New York Times

Can we predict County
Election Outcome
given County's Socio-
Economic Data?

*How Data Failed Us
in Calling an Election*



Michigan Democratic Party members in Flint studying precinct results on Tuesday. Virtually all major vote forecasters put Hillary Clinton's chances of winning in the range of 70 to 99 percent.
Brittany Greeson for The New York Times

Data Overview

County Data includes the following information (categories):



- Education



- Healthcare



- Occupation



- Economic



- Demographic



- Crime

Data Overview

County Data includes the following information (categories):



- Education



- Healthcare



- Occupation



- Economic



- Demographic



- Crime

OUTCOME

ACTUAL



Won 486 counties

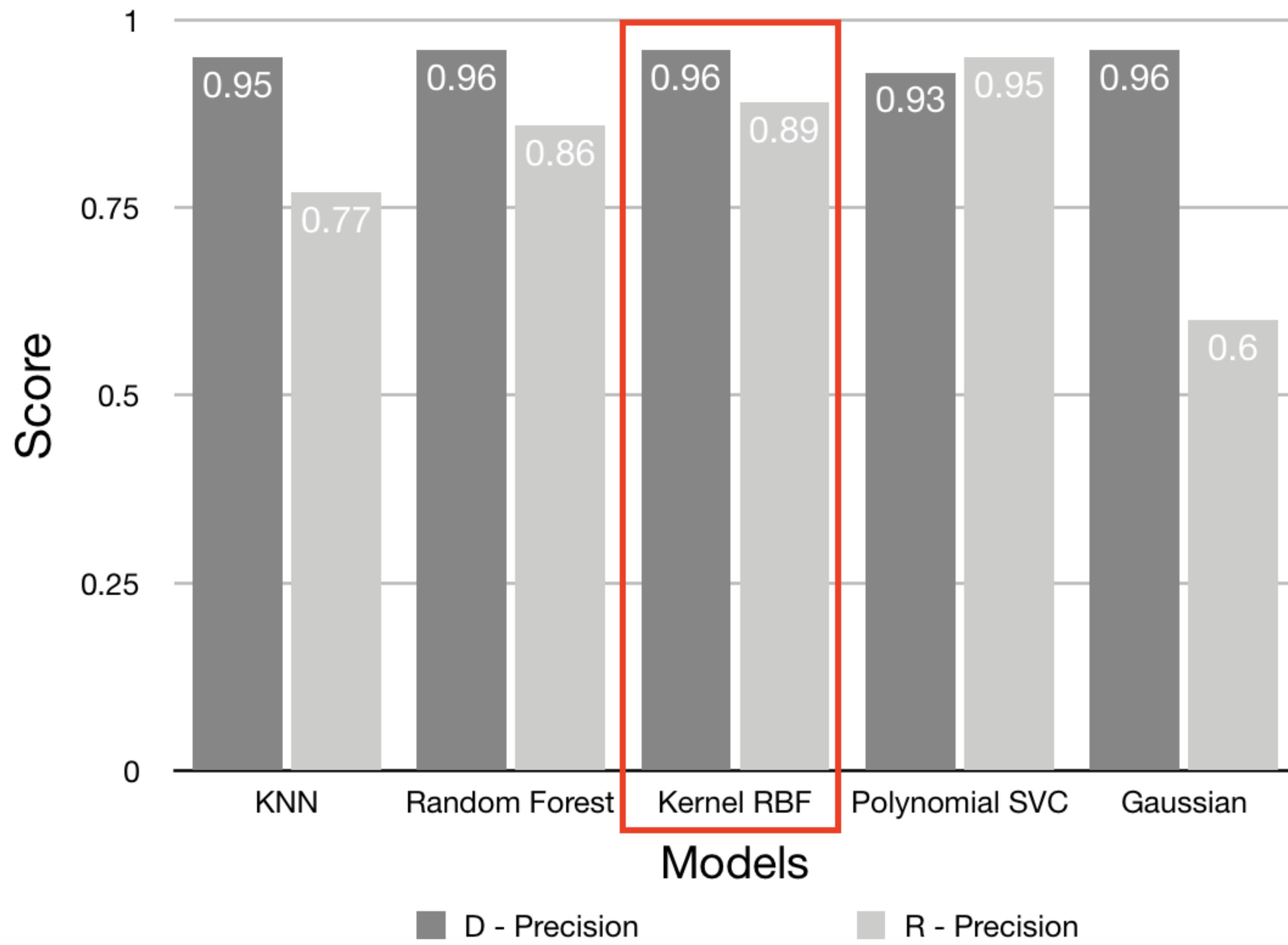


Won 2625 counties

Data Modeling

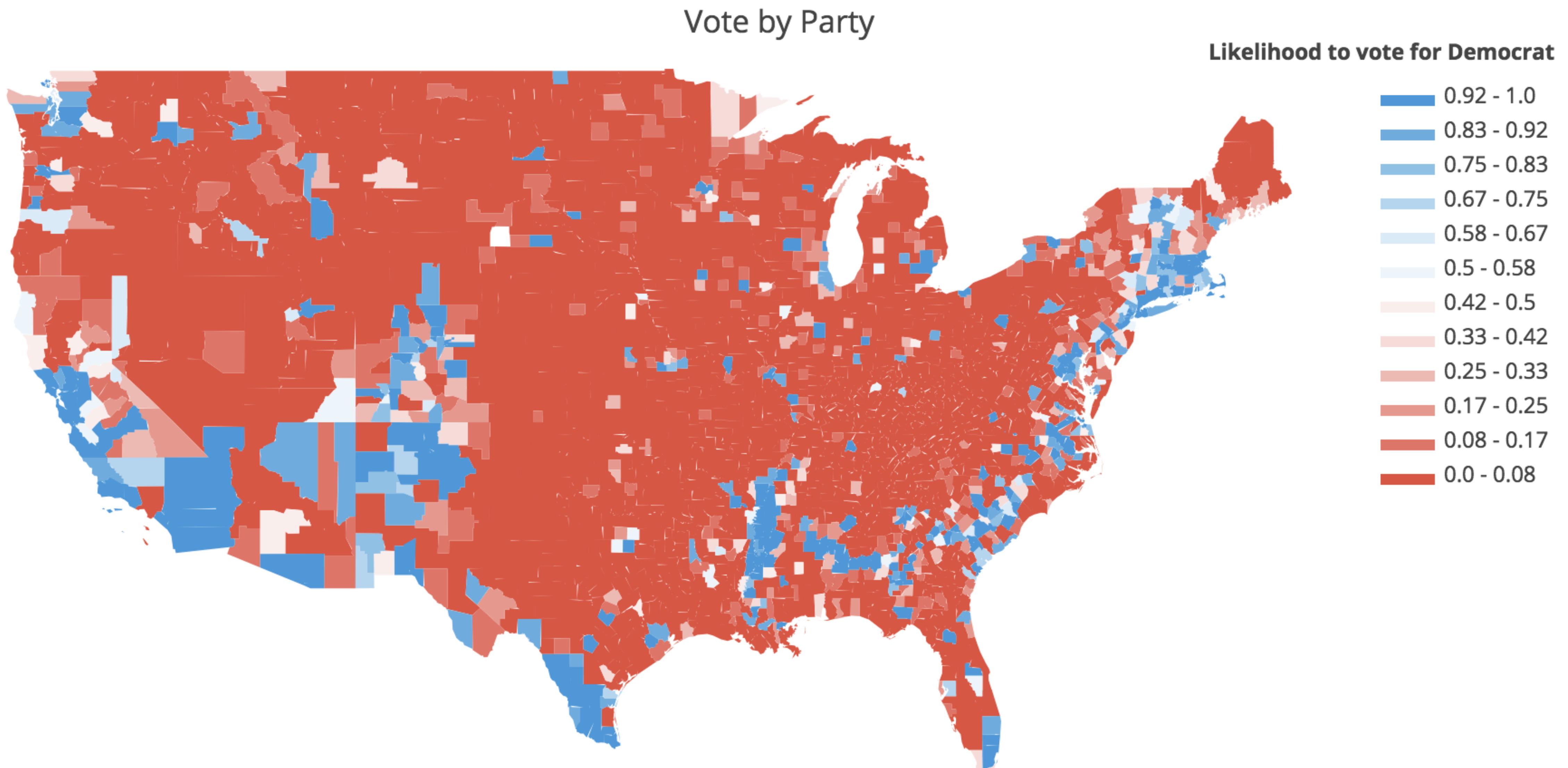
$$\text{Precision} = \frac{tp}{tp + fp}$$

Model Comparison

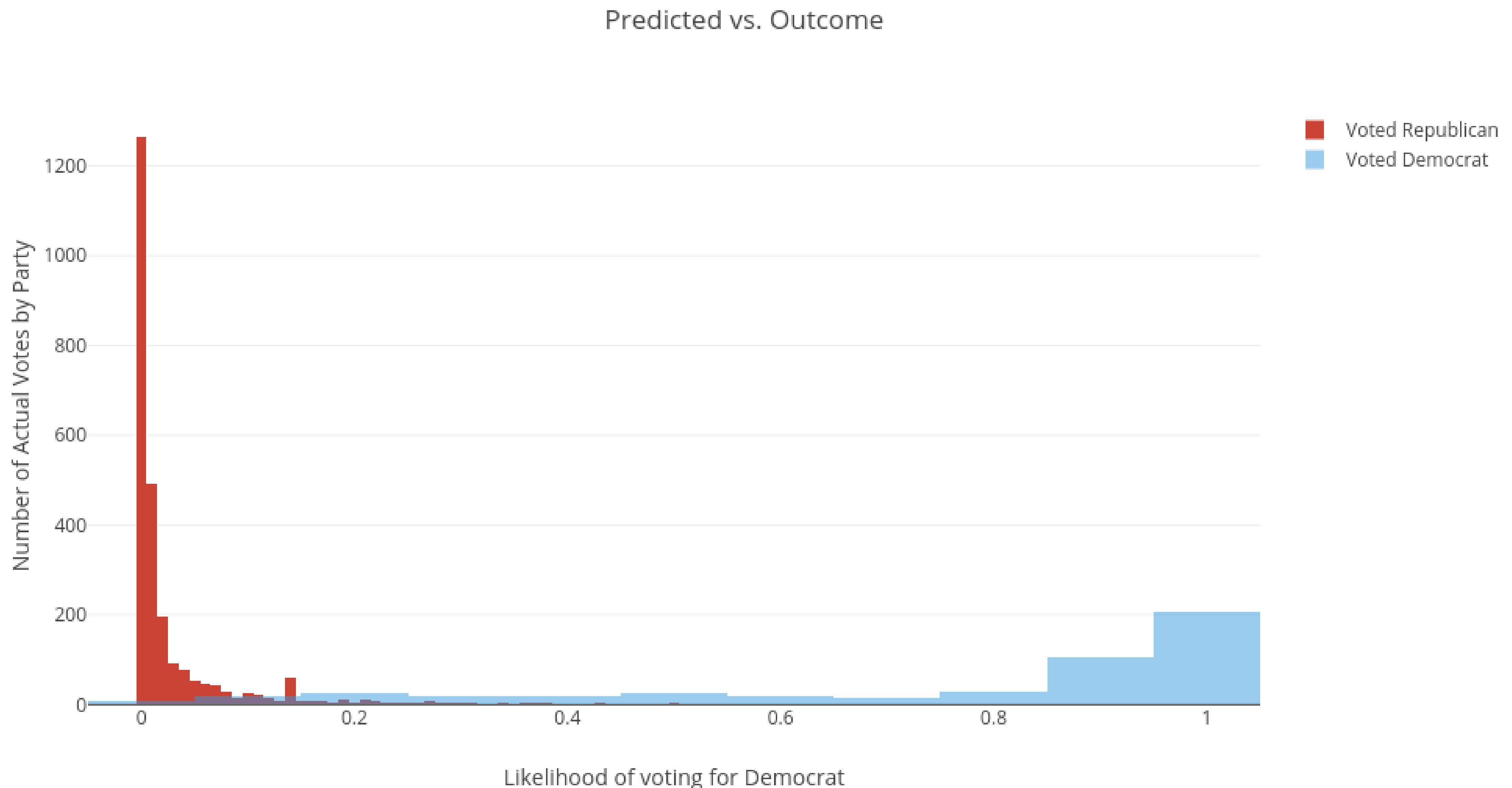


Best Model is Kernel RBF [$C=80$, $\gamma = .001$]

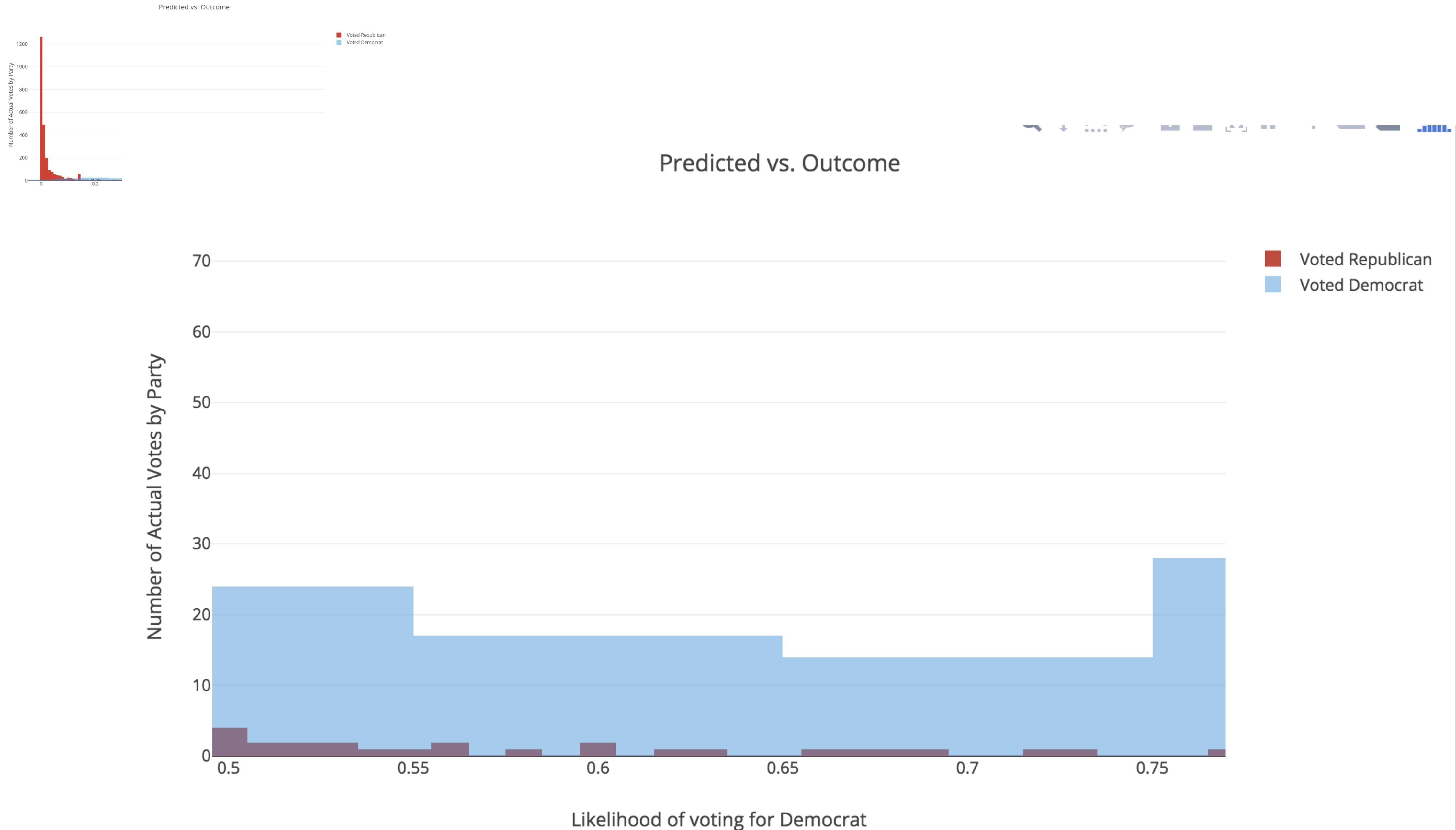
Predicted Outcome



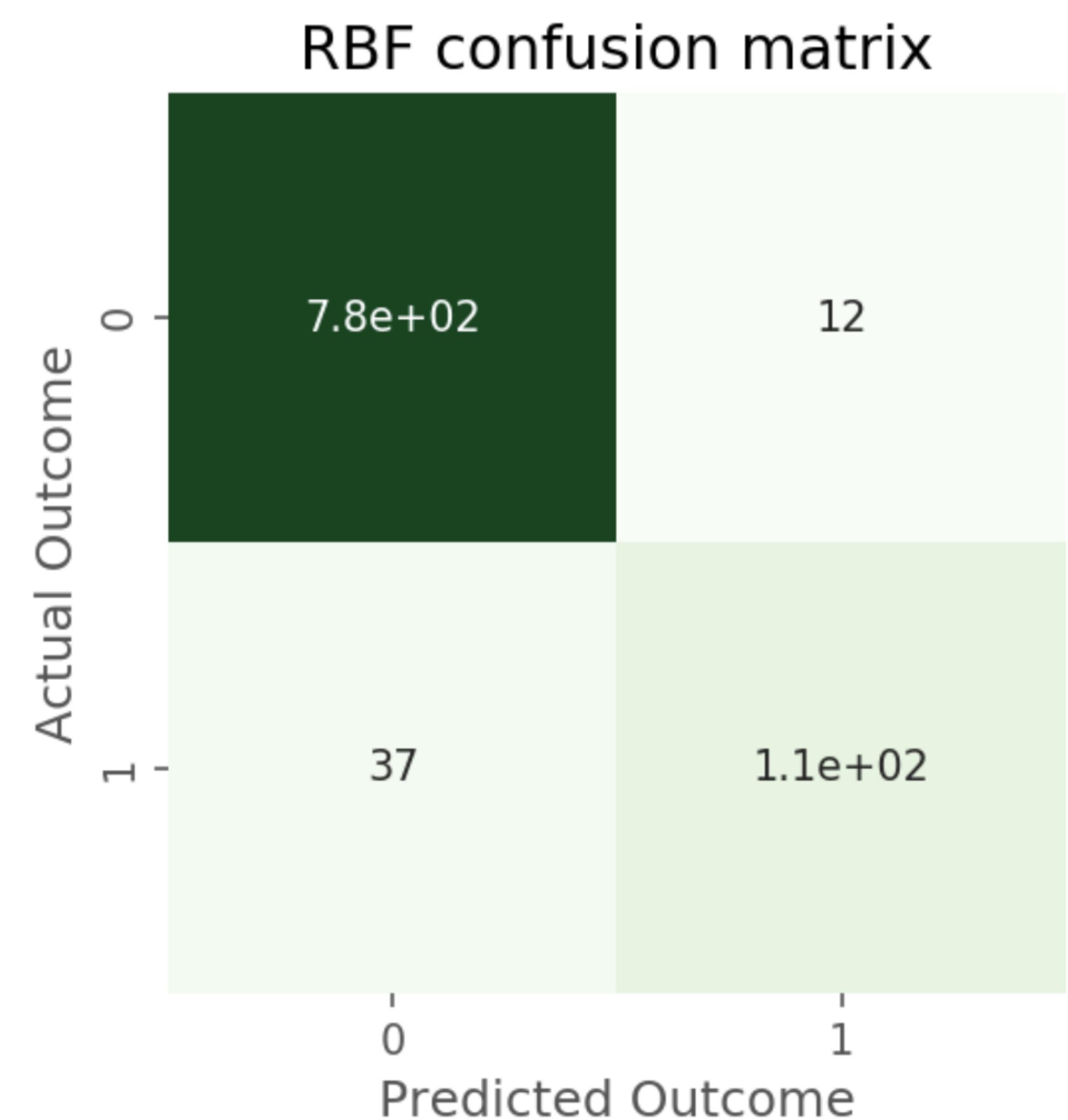
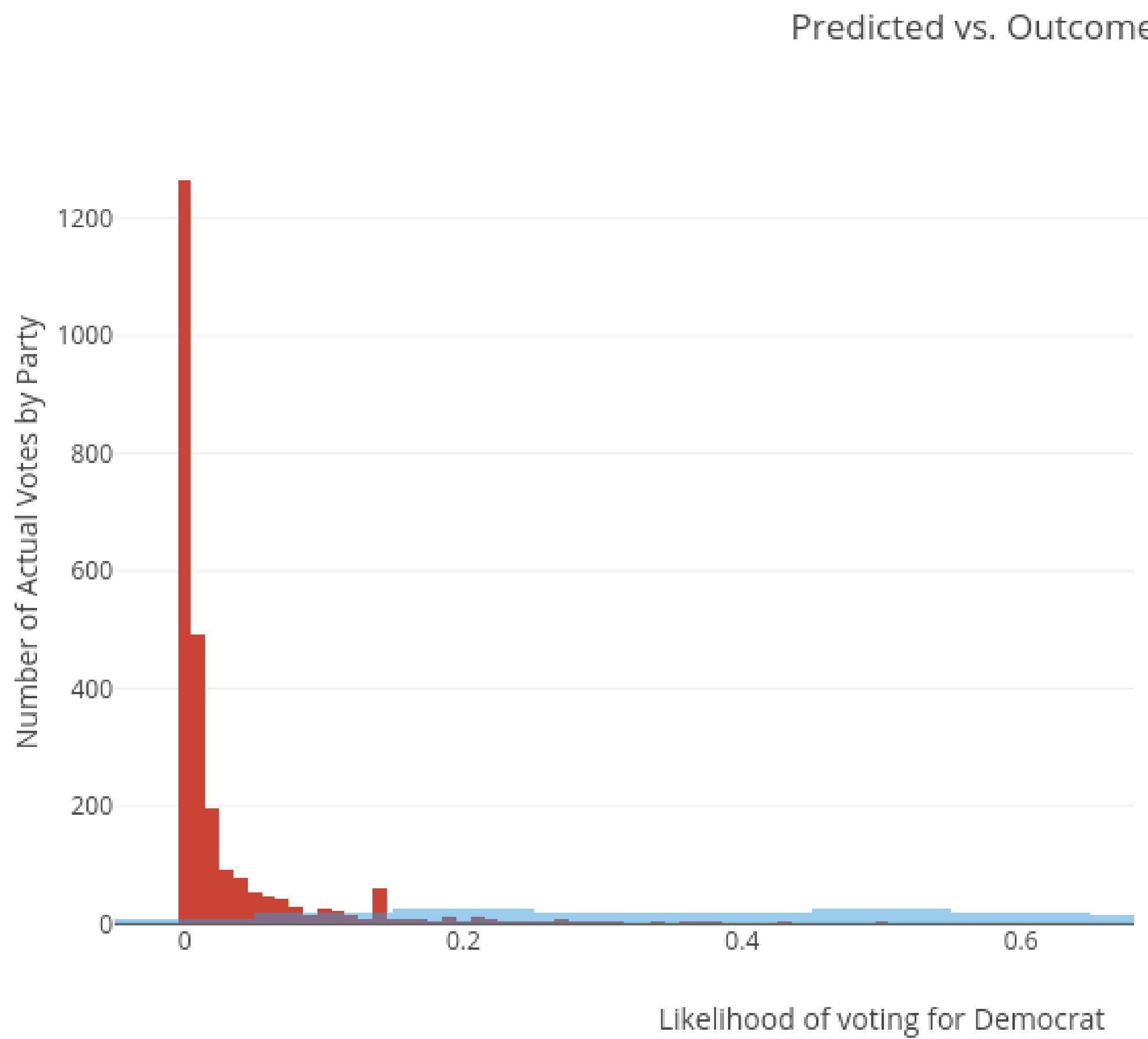
Predicted vs. Actual



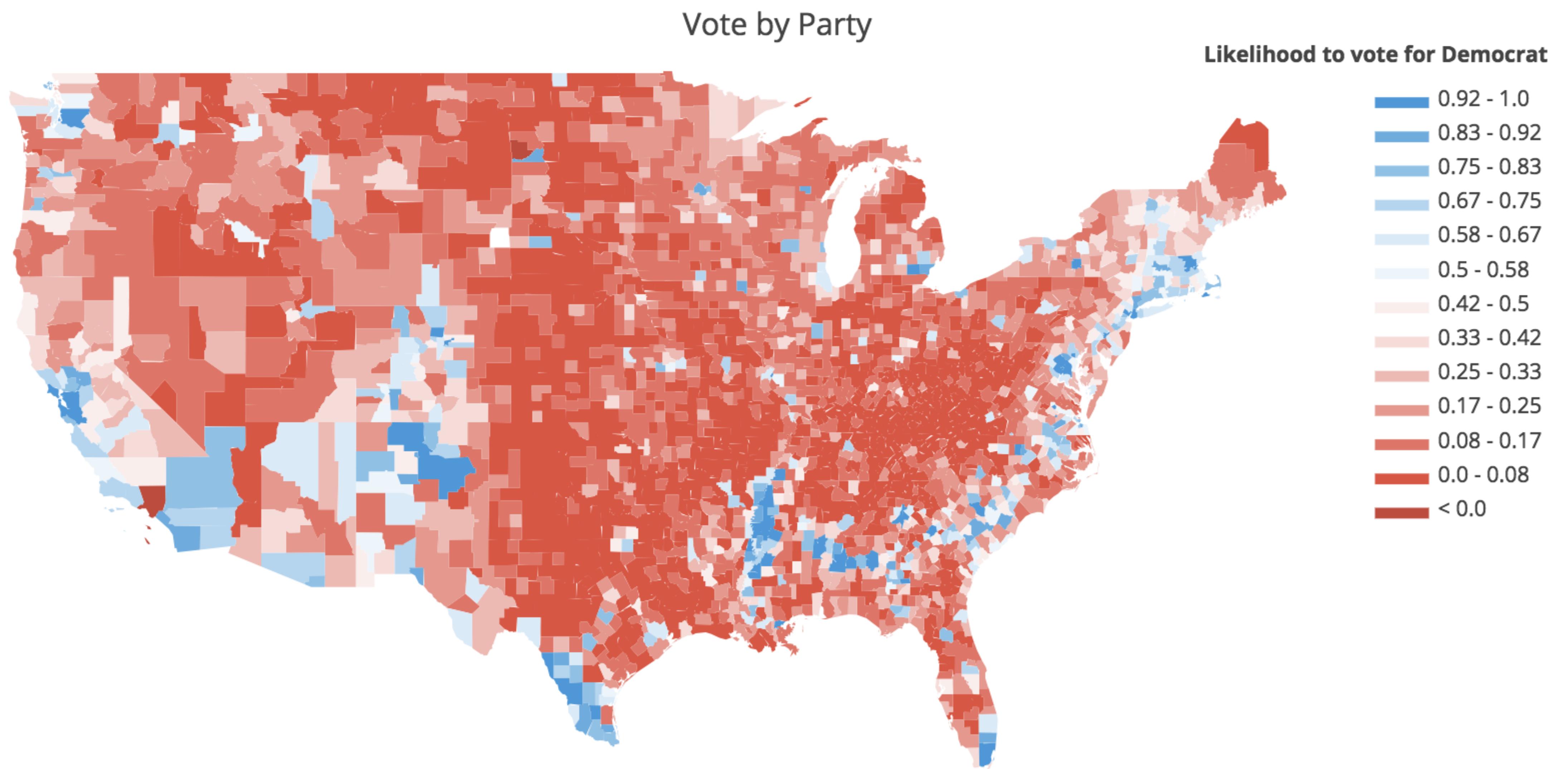
Predicted vs. Actual



Predicted vs. Actual



Monte Carlo



Next Step

Replicate Dave Wasserman's analysis to prove its truthfulness (if true).
Obtain Healthcare Information from Institute for Health Metrics and Evaluation and perform analysis.

 **Dave Wasserman** 
@Redistrict  ▾

Donald Trump won 76% of counties w/ a Cracker Barrel & 22% of counties w/ a Whole Foods -- a 54% gap. In '92, gap b/t same counties was 19%.

10:52 AM - 9 Nov 2016

1,701 Retweets 1,958 Likes 

127 1.7K 2.0K

The presidential election
Illness as indicator

Local health outcomes predict Trumpward swings

Nov 19th 2016 | NEW YORK



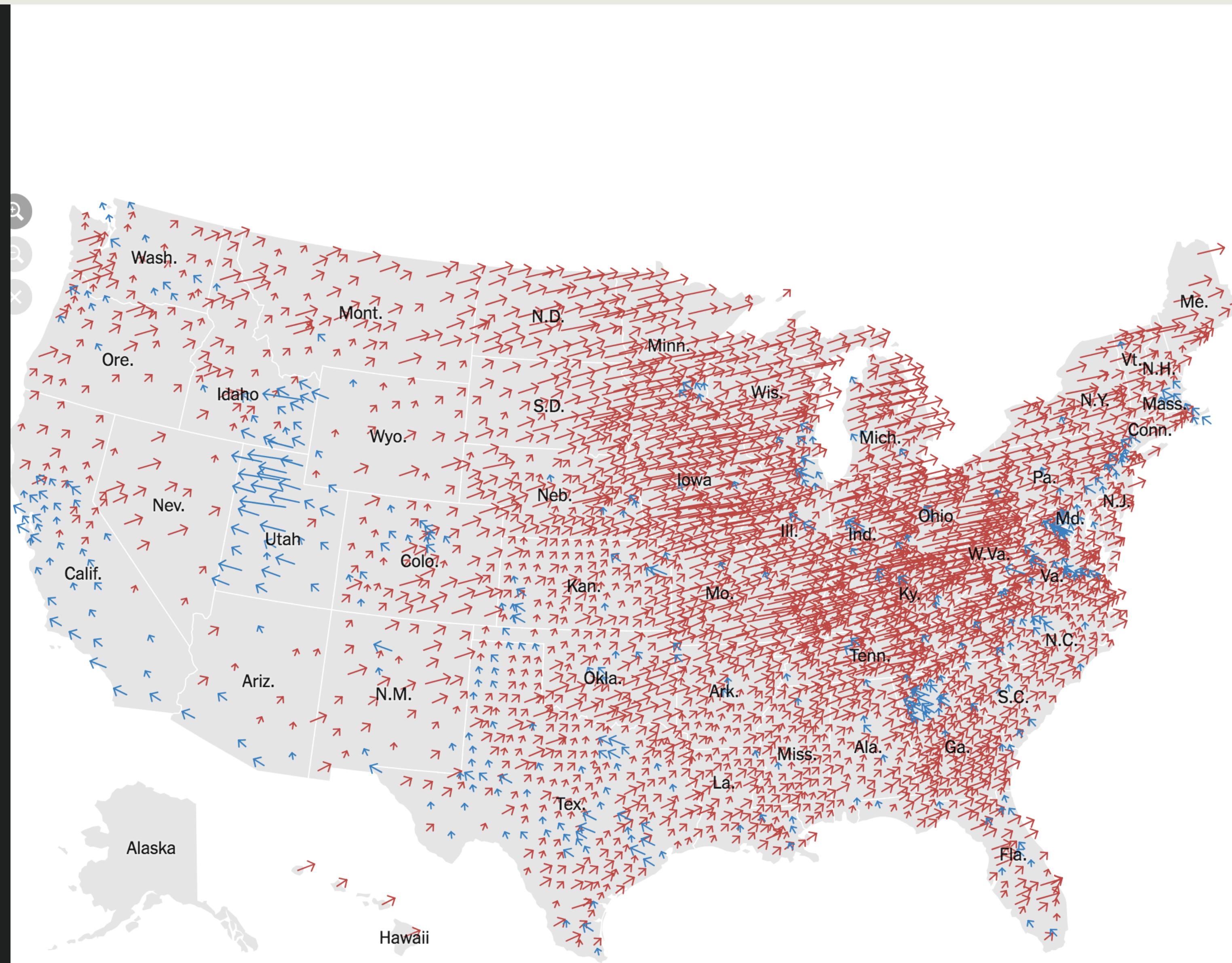
Like 89

Tweet

THE first piece of news Americans woke up to on November 9th was that Donald Trump had been elected president. The second was that he owed his victory to a massive swing towards Republicans by white voters without college degrees across the north of the country, who delivered him the rustbelt states of Michigan, Wisconsin and Pennsylvania—

Conclusion

As Data Scientists, we should have a full grasp of our data limitations and the potentially flawed assumptions that we make to build models as to better adjust and understand the applicability.



Likelihood of "Wrongness"

