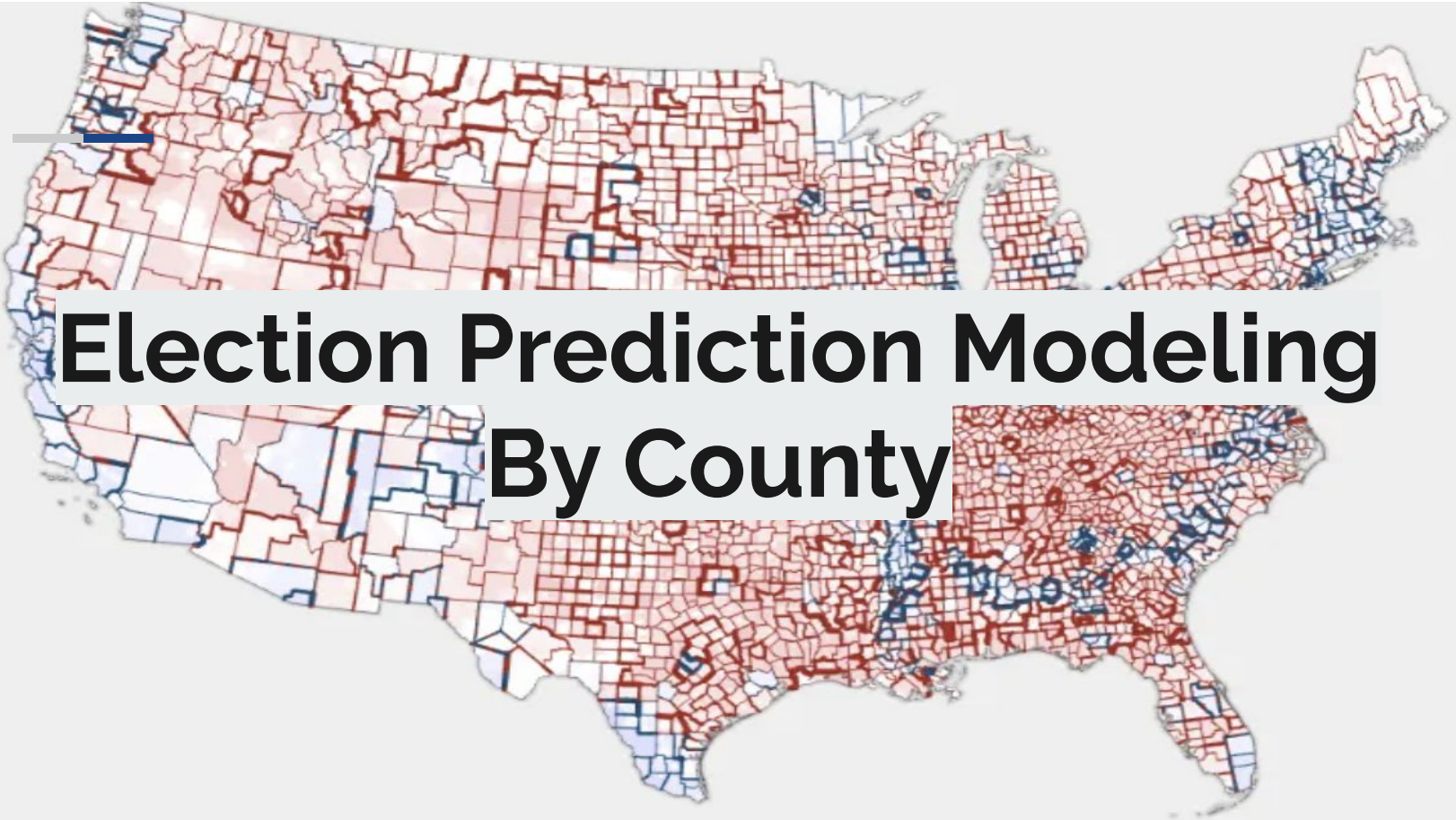


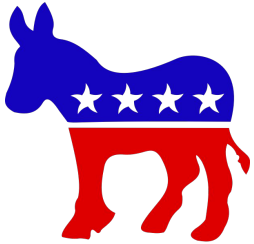
Election Prediction Modeling By County

Agenda

1. Overview
2. The Approach
3. Recommendations
4. Next Steps
5. Conclusion

Overview

- The Democratic Party faces an uncertain political future.



- EMG Consulting has been retained to build a model to predict election outcomes to help the party better utilize its resources.

EMG CONSULTING
DATA. ANALYSIS. INSIGHTS.

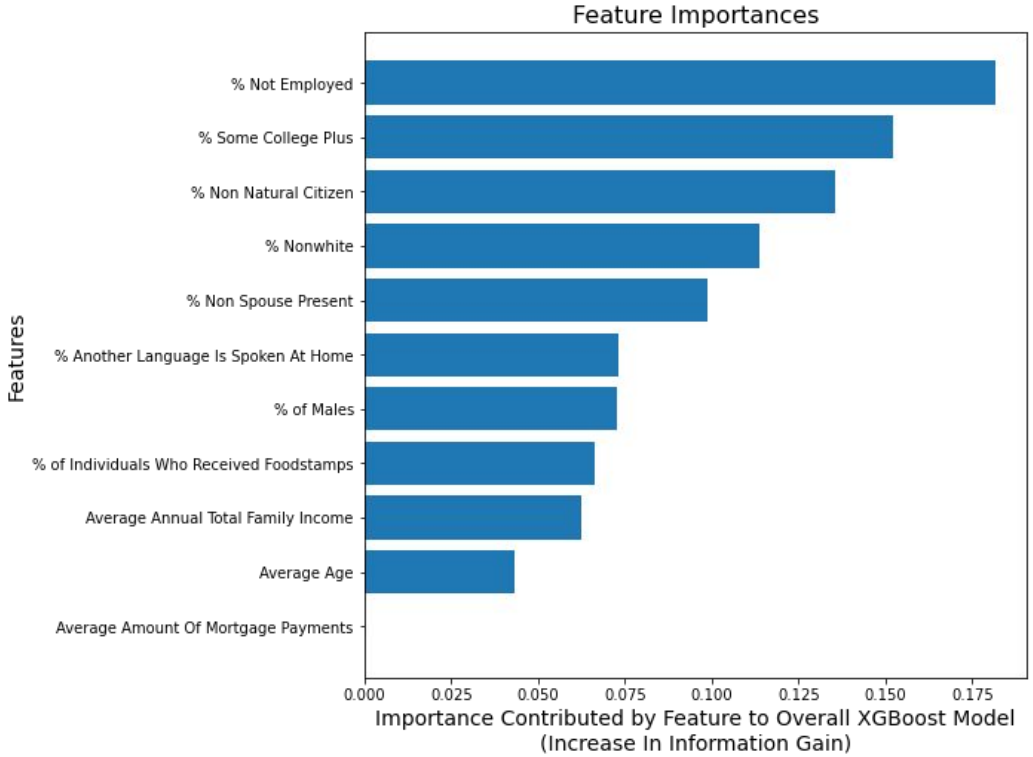
The Approach

- Roughly 1,600 records and 27 columns of county-level demographic and presidential election data from 2008-2020 — approx. 50/50 Democratic/Republican split
- Make predictions based on similarity to counties with known outcome in these elections
- Emphasis on creating a model that makes few predictions where the Republican victory comes as a surprise — final model correctly identified about 80% of Republican victories

Recommendation #1: Keep a particularly careful eye on counties where a large proportion is not in the labor force.

Proportion of the population out of the labor force is important.

- Includes retirees - a key Republican constituency
- In a different model, increase of ~6.5% percent in this proportion of the population was associated with a ~28% increase in likelihood of Republican victory



Recommendation #2: To prioritize avoiding surprising election losses, use a model that classifies counties based on similarity to counties with known outcomes.

We have a few important metrics:

Accuracy

How often are we right?

Total proportion of correct predictions.

Recall

How often are we surprised in a bad way?

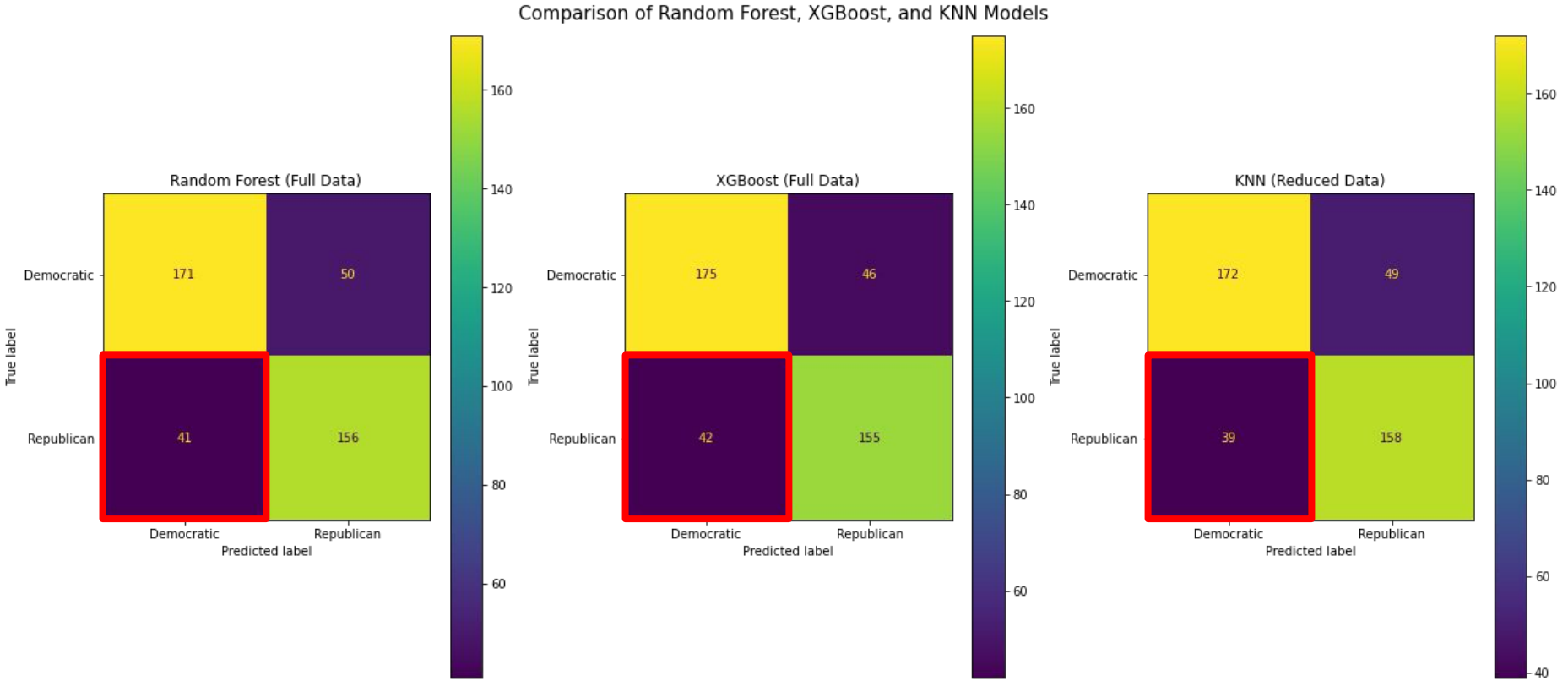
False negatives mean surprise Republican victories.

Precision

How often are we surprised in a good way?

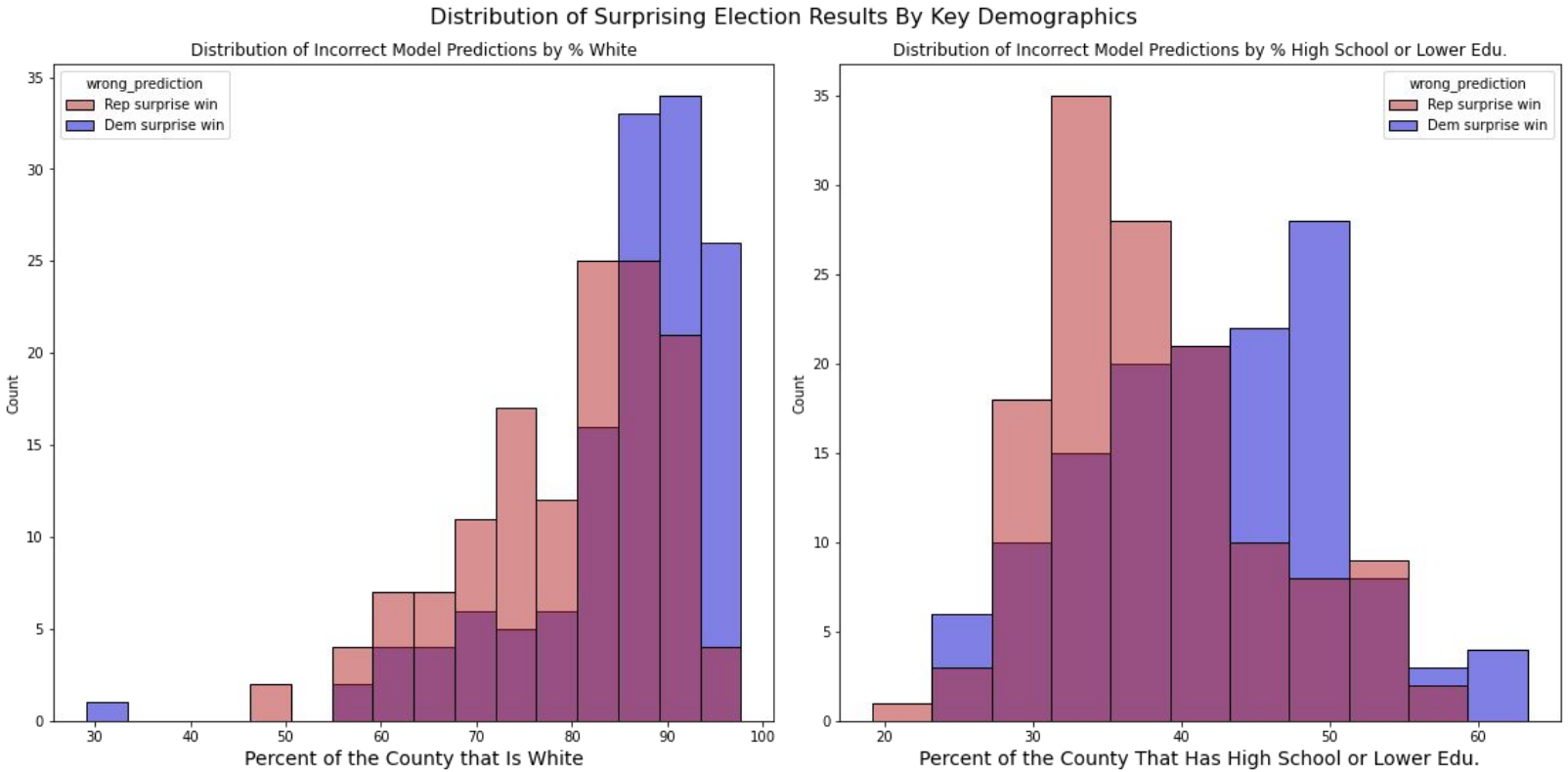
False positives mean surprise Democratic victories.

Model Performance



Recommendation #3: Work to shore up the party's base in communities of color and higher-educated communities.

The Map Is Shifting. The Party Must Adapt.



Next Steps

- Include more counties in the dataset from the last four elections to increase the amount of training data available.

- Examine how the model's incorrect predictions may have differed by year to see if there are any trends in surprising results (e.g. Republican victory years vs. Democratic victory years).

- Evaluate how county vote shares and demographics have changed over time to see if any relationships exist.

Thank you.

Contact:
emgerber94@gmail.com

GitHub Repository Link