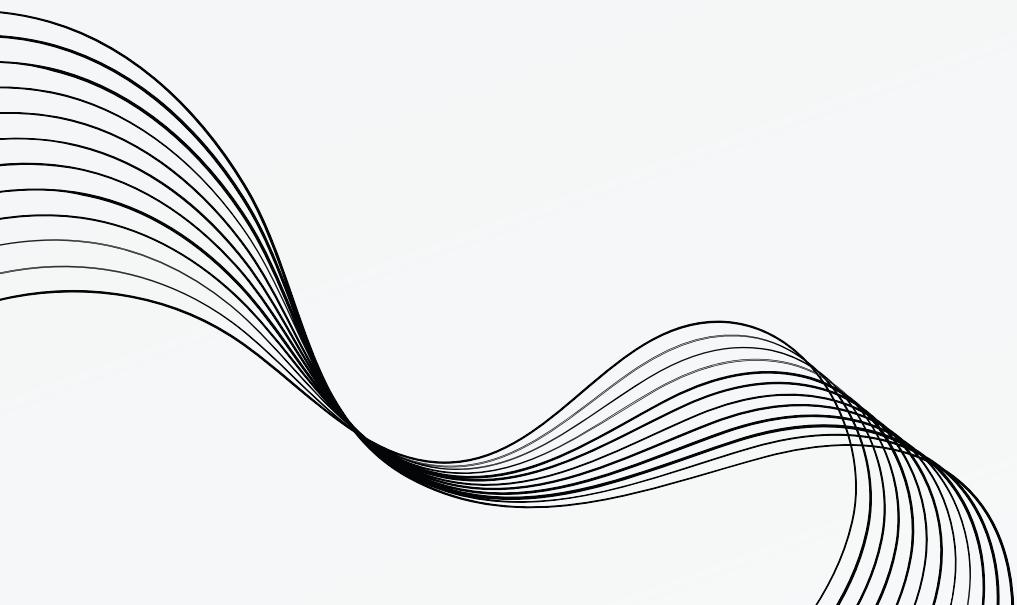


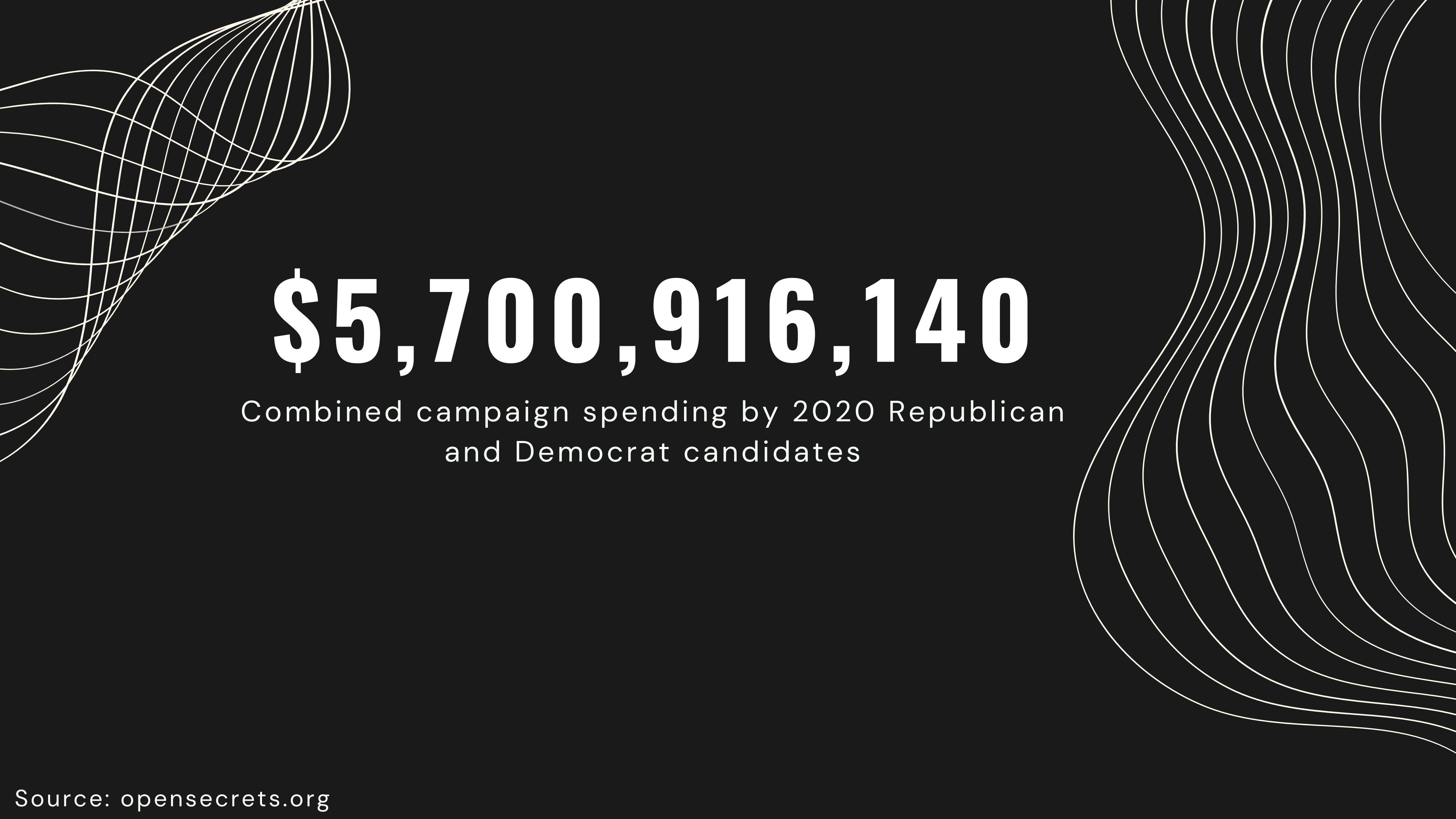


POLITICAL

WHAT IS POLITICALL

Politicall is a machine learning program that aims to predict whether a county will vote Republican or Democrat based on the input of three social vulnerability factors.





\$5,700,916,140

Combined campaign spending by 2020 Republican
and Democrat candidates

ANALYZED DATA

Social vulnerability index

Dataset containing 3143 rows and 158 columns.

Procured from the Agency for Toxic Substances and Disease Registry.

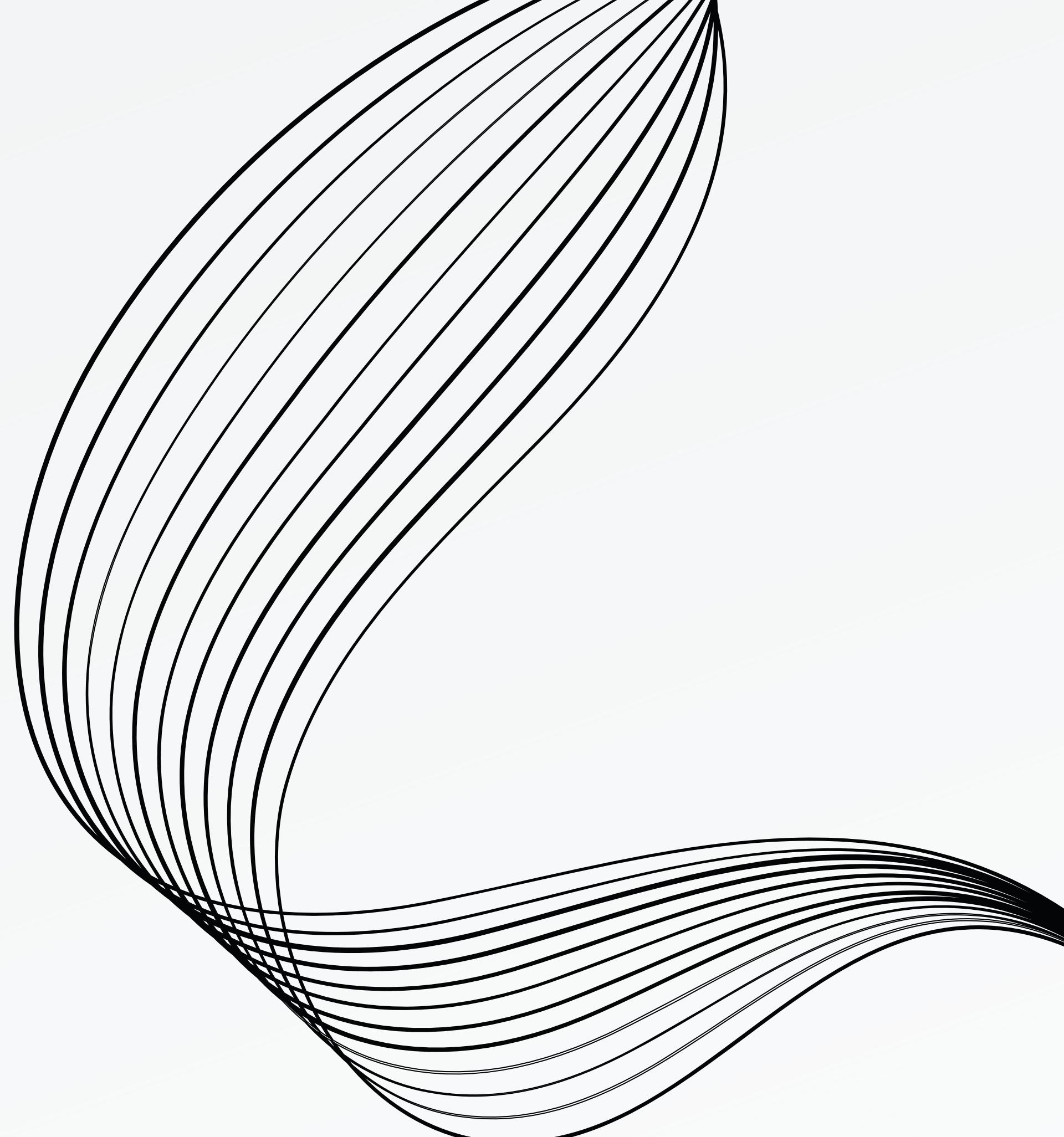
2000–2020 elections by U.S. county

The dataset contains 72,619 rows and 12 columns.

Procured from the Harvard Dataverse.

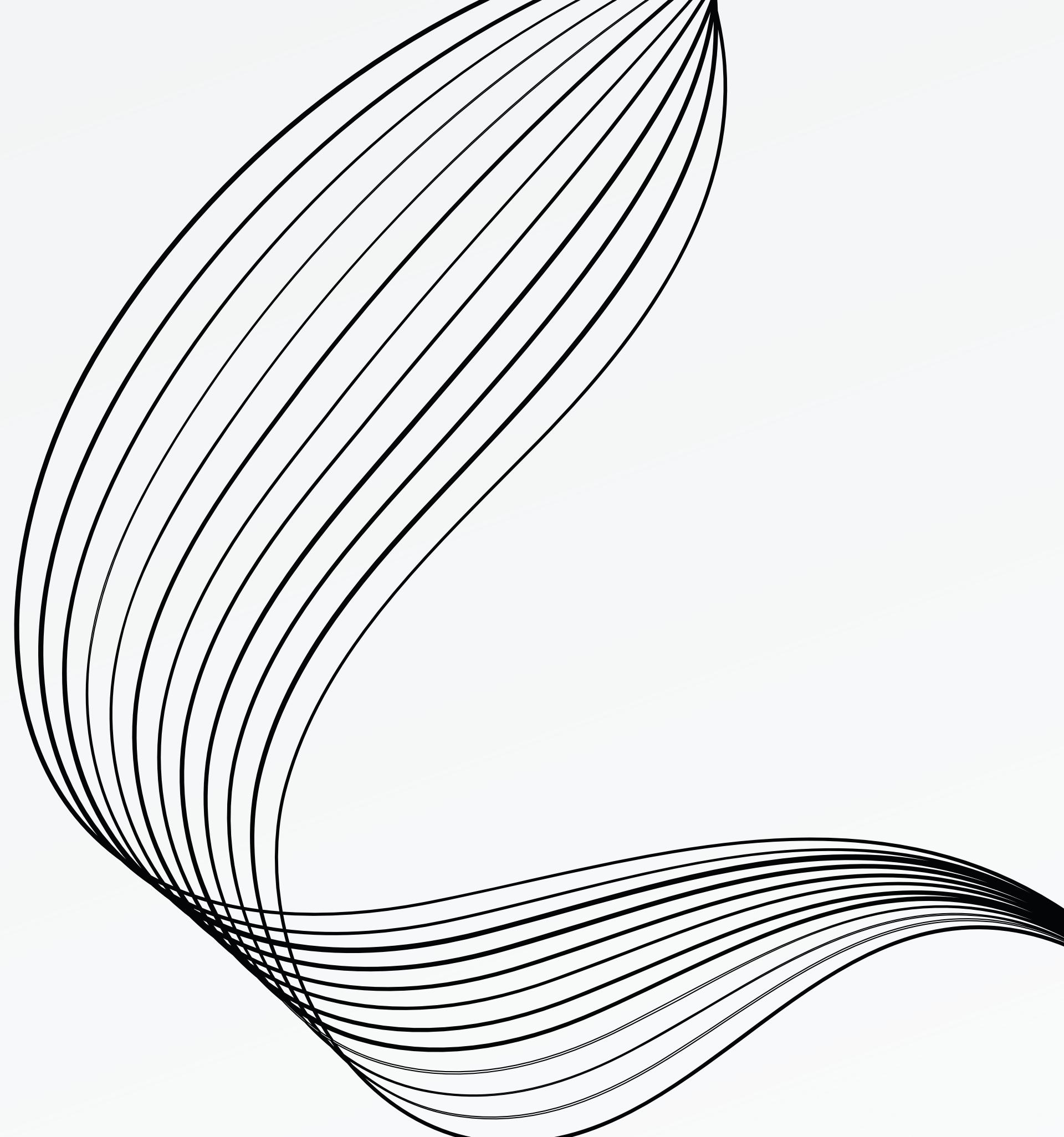
CLEANING

- Checked for null values
- Dropped all but estimation columns from SVI dataset
- Dropped all but 2020 election rows from election dataset
- Dropped all but the winning candidate per county from election dataset
- Renamed counties with different spellings in election dataset
- Dropped counties where missing from either dataset



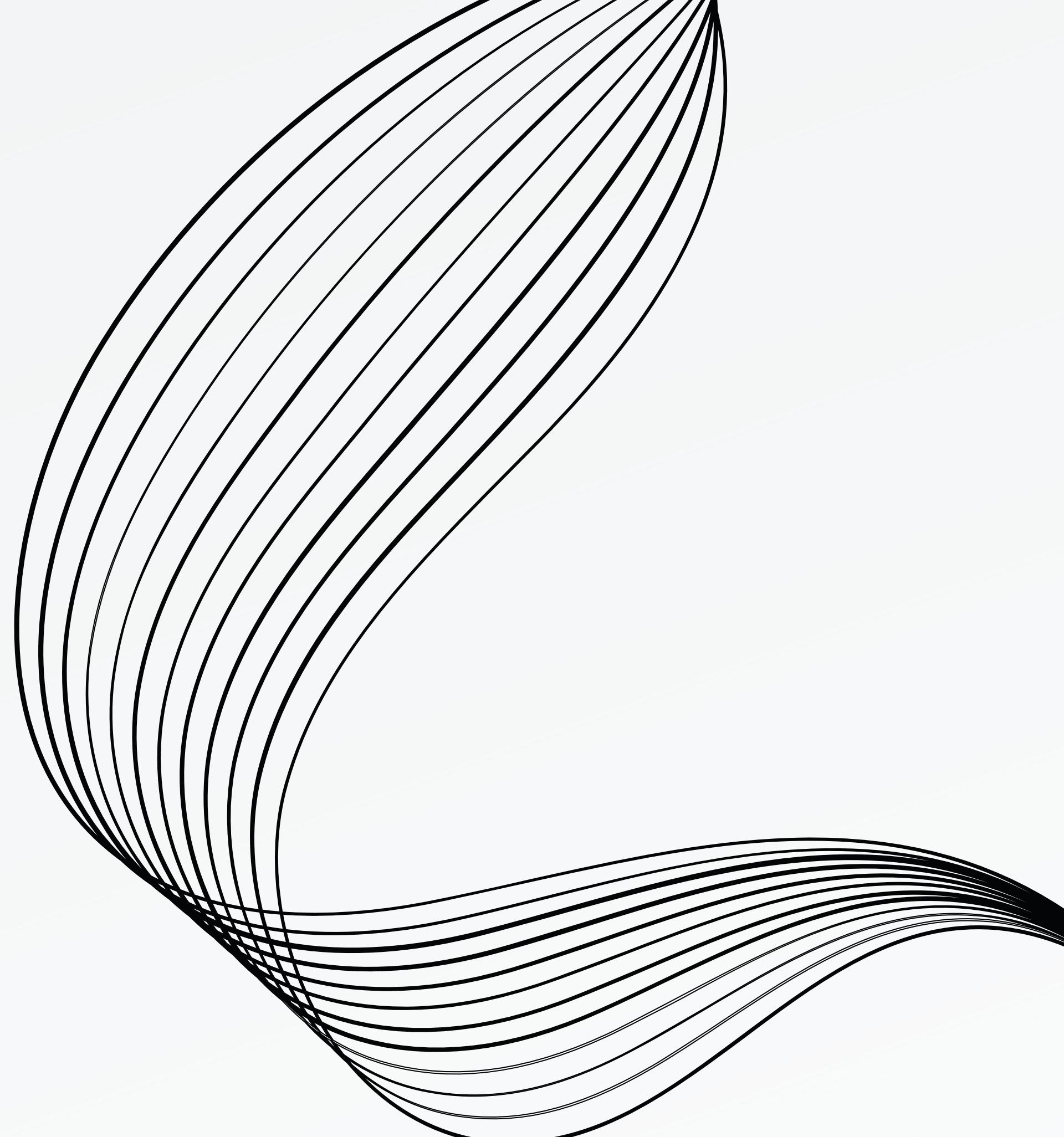
PRE- PROCESSING

- Merged datasets
- Defined 'PARTY' as target variable
- Split data
- Transformed data using standard scaler
- Instantiated a logistic regression model
- Fit training data to model



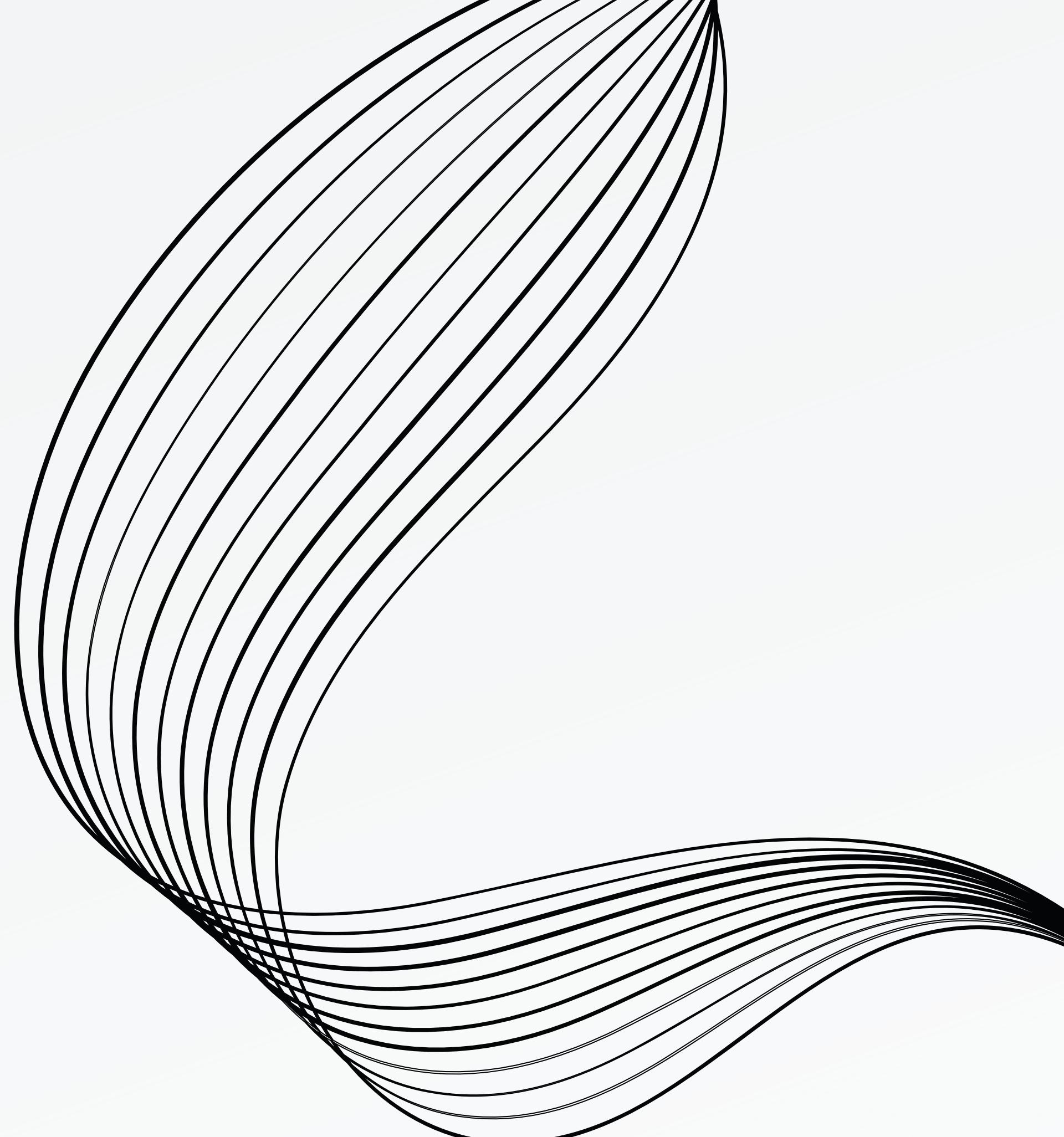
MODEL EVALUTION

- Tested the logistic regression model, looking for accuracy
- Cross-validated with other models
- Checked for overfitting with best model



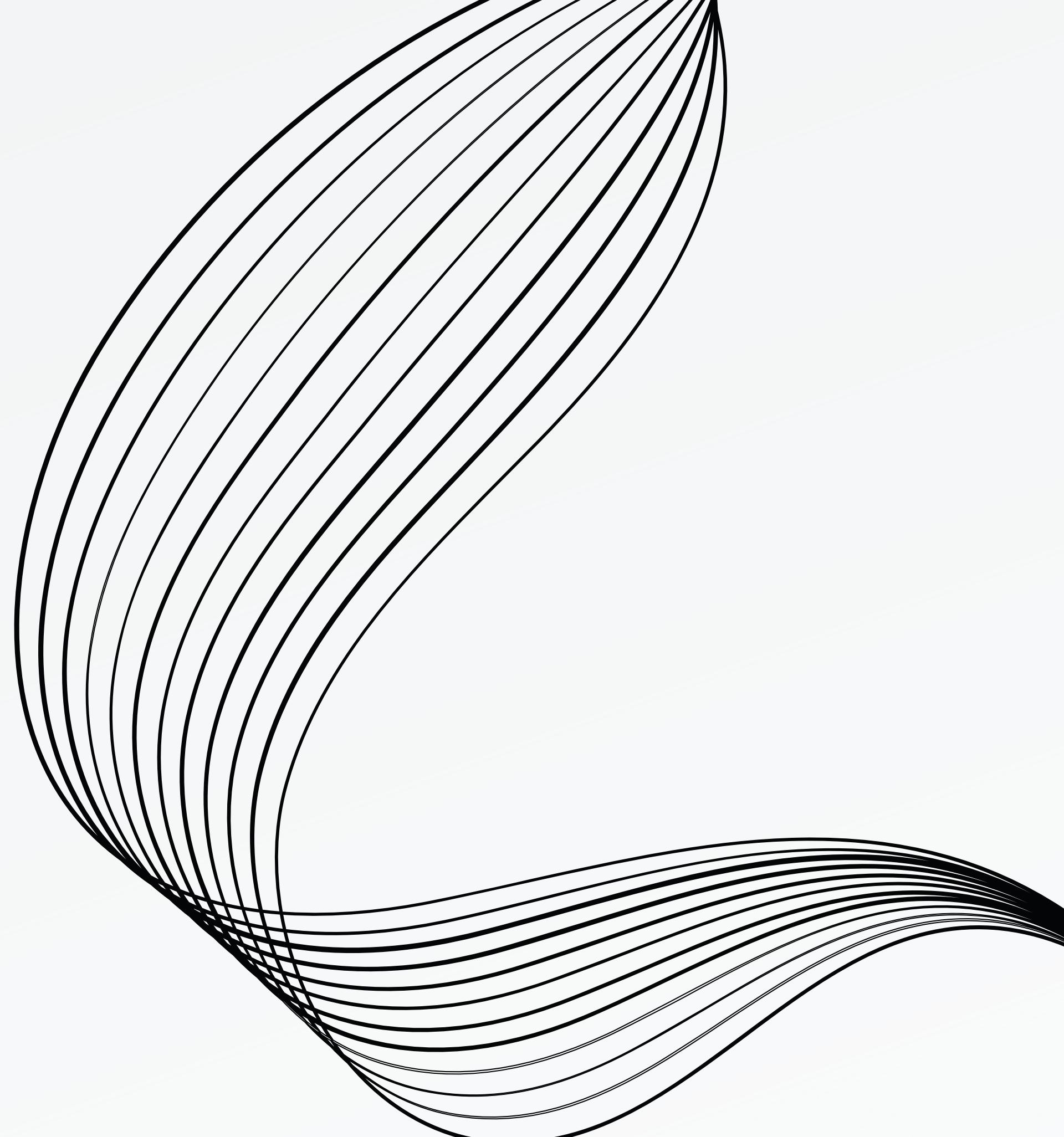
FEATURE SELECTION

- Ran coefficient test
- Reduced to three features:
 - number of non-caucasians;
 - number of buildings with 10 or more housing units;
 - number of people with no vehicle.
- Reran cross validation
- Checked for overfitting on best model



BALANCING DATASET

- Upsampled dataset using SMOTE
 - `y_train.value_counts` from:
 - REPUBLICAN: 2067
 - DEMOCRAT: 422
 - `y_train.value_counts` to:
 - REPUBLICAN: 2067
 - DEMOCRAT: 2067
- Scaled upsampled data

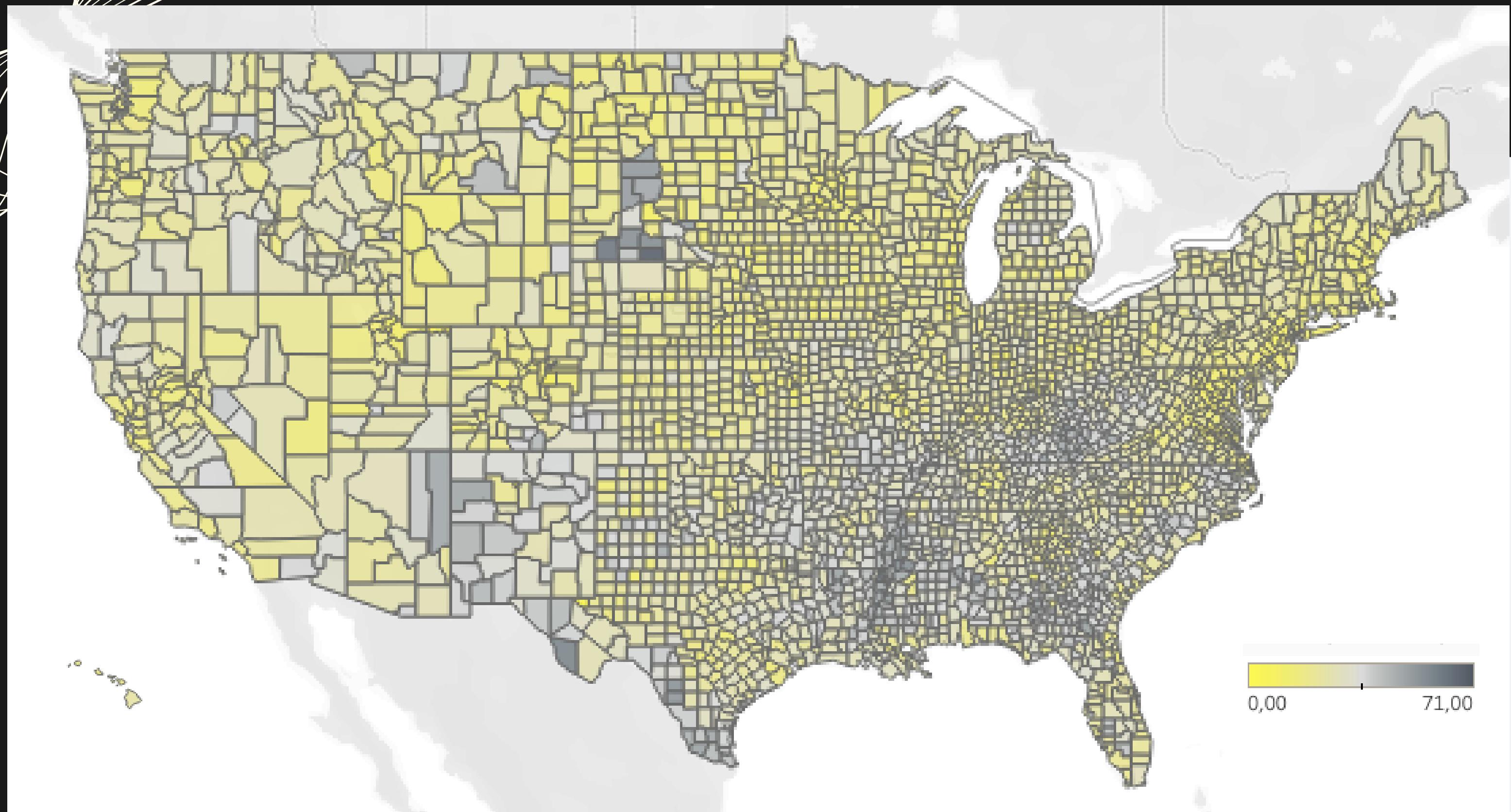


FINAL RESULT

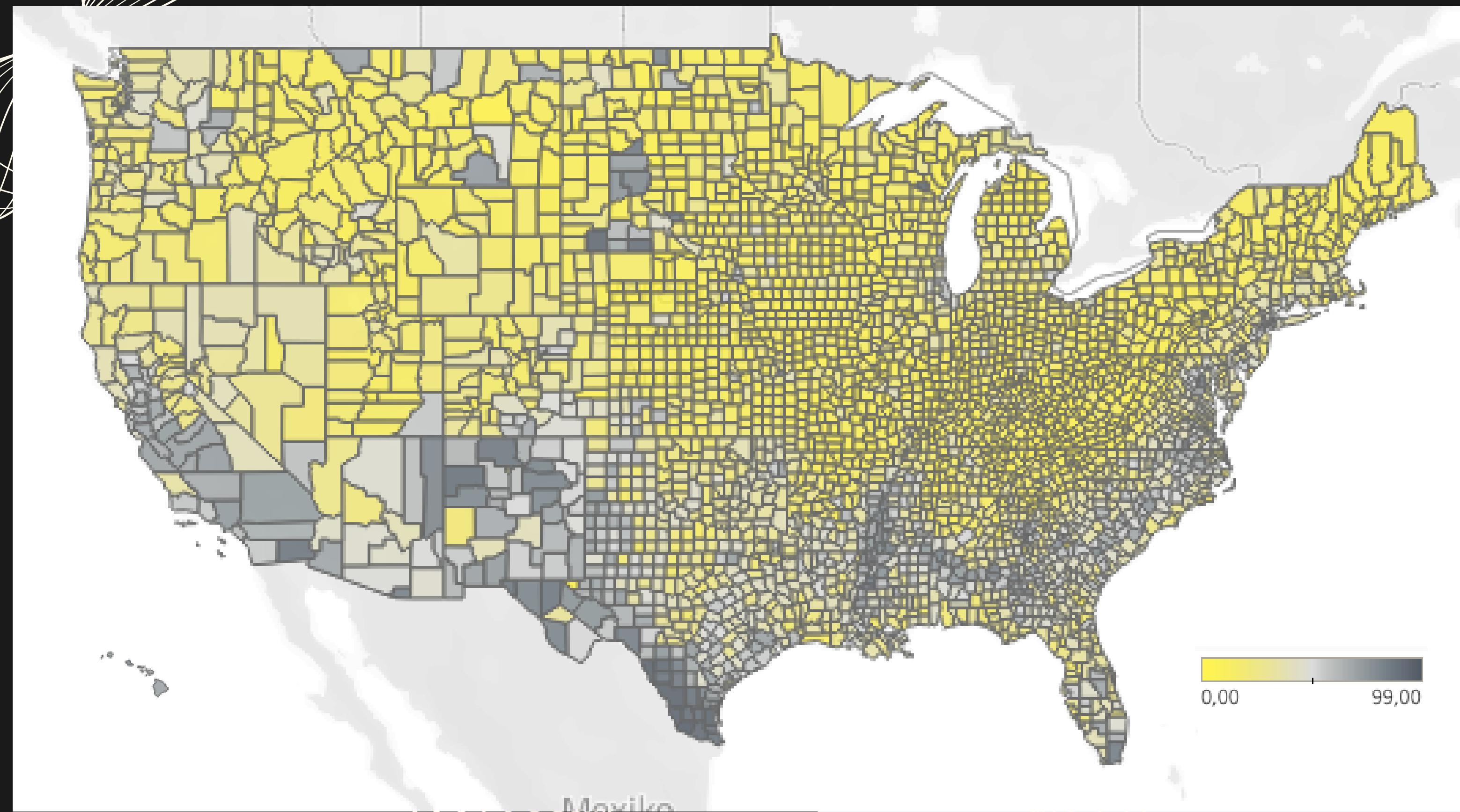
- Logistic regression model
- Accuracy: 82.98



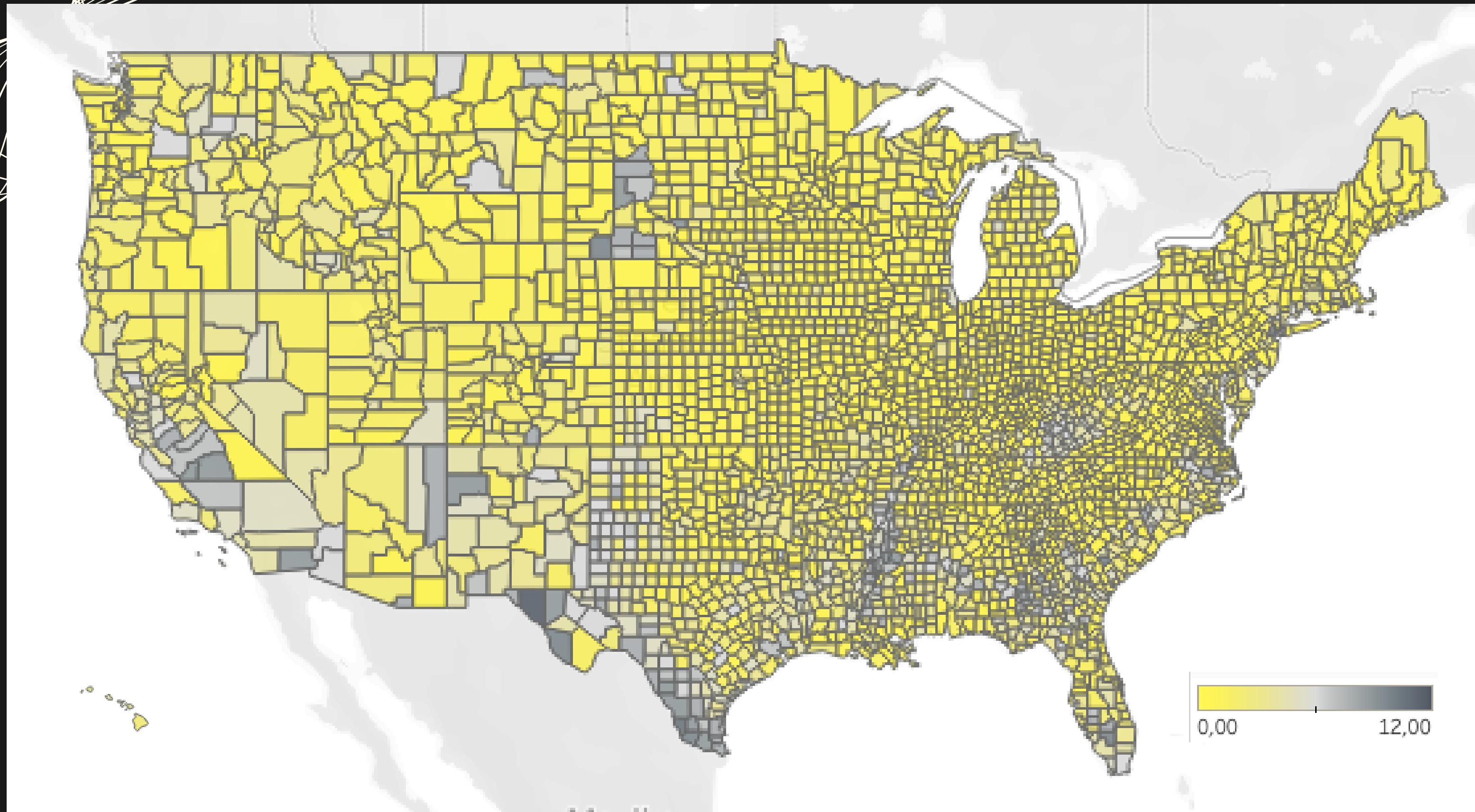
PERCENTAGE OF 150% BELOW POVERTY



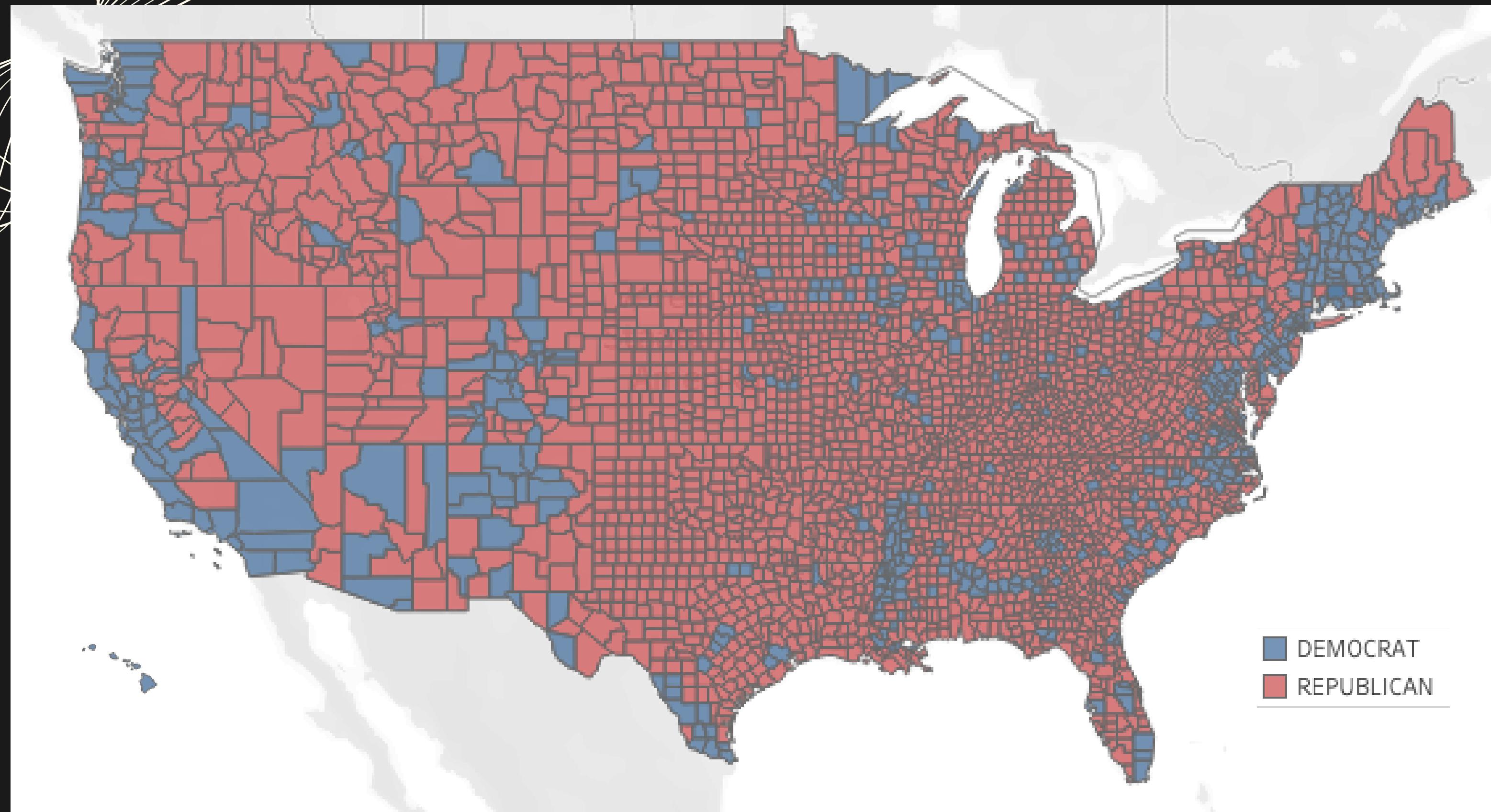
PERCENTAGE OF NON-CAUCASIANS

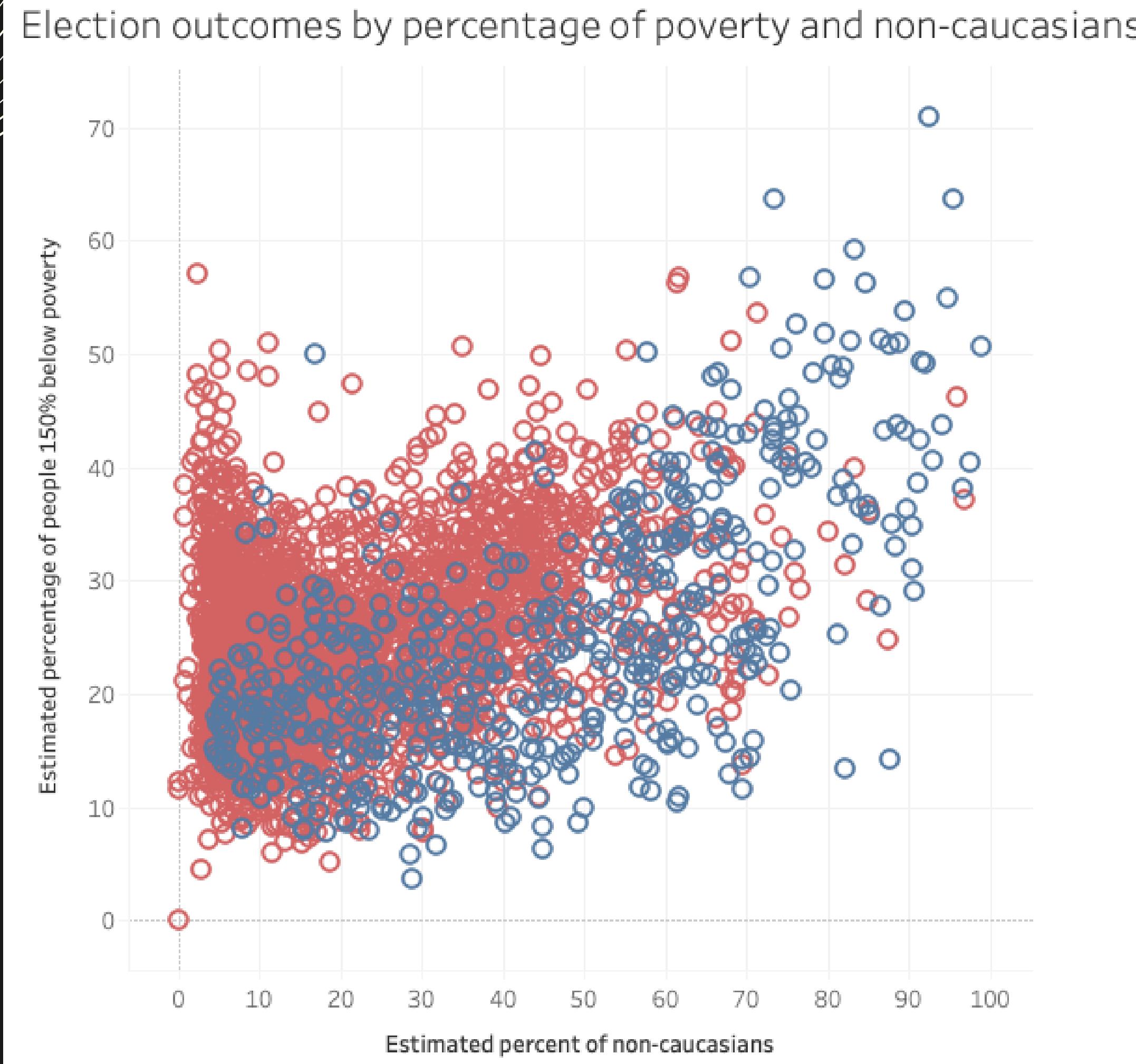


NUMBER OF 90% PERCENTILE SVI FLAGS

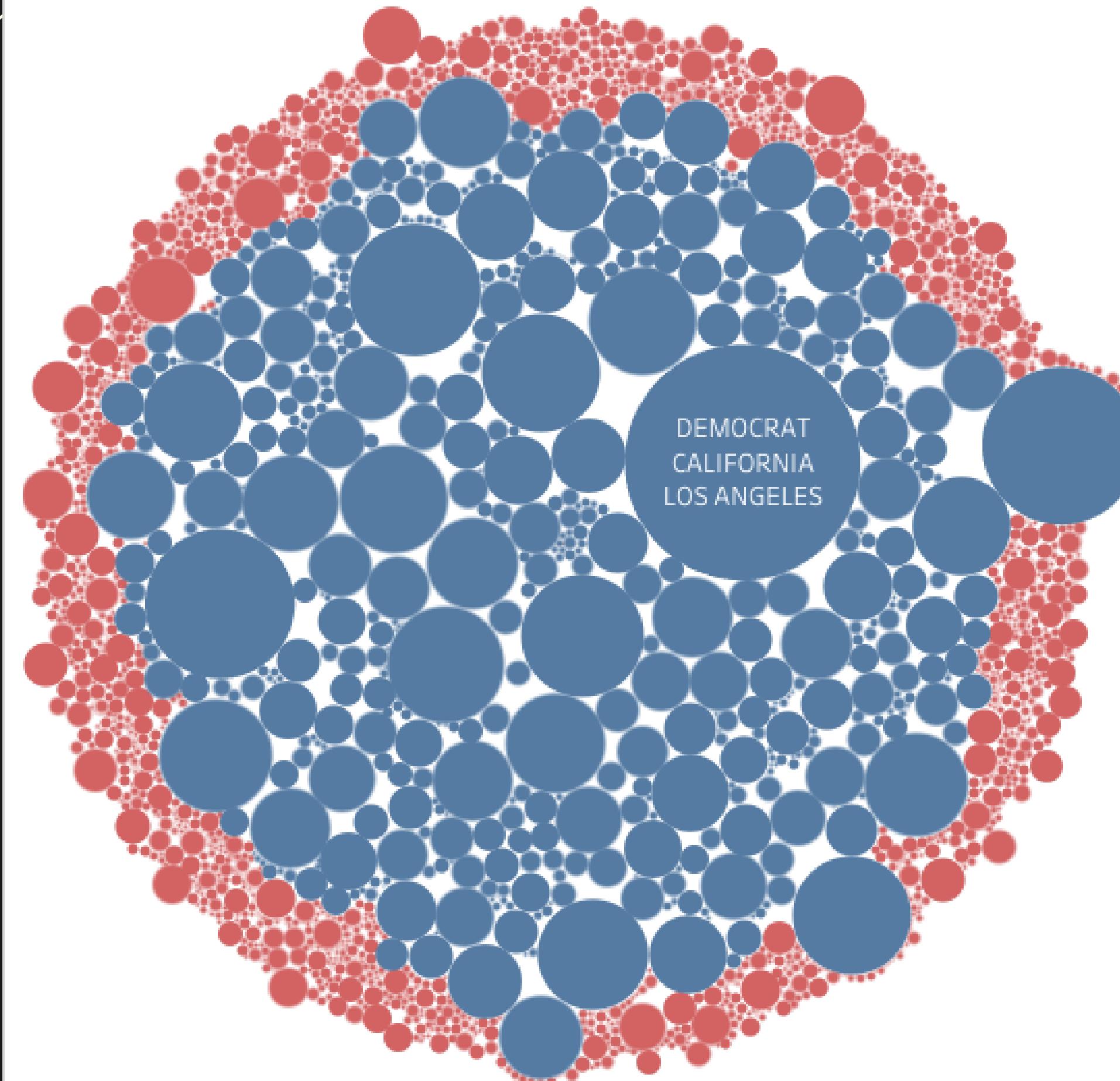


ELECTION OUTCOMES BY COUNTY

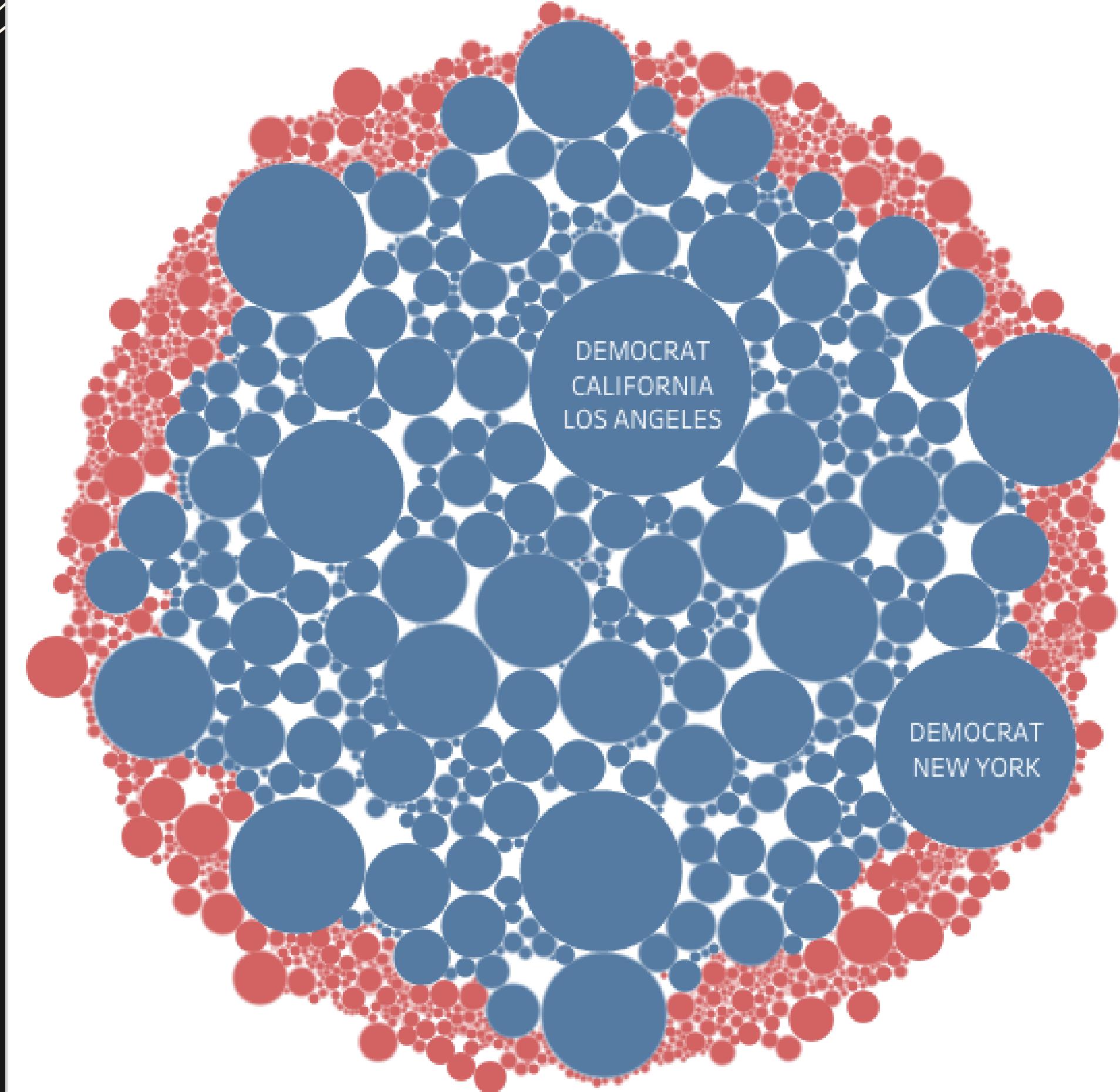




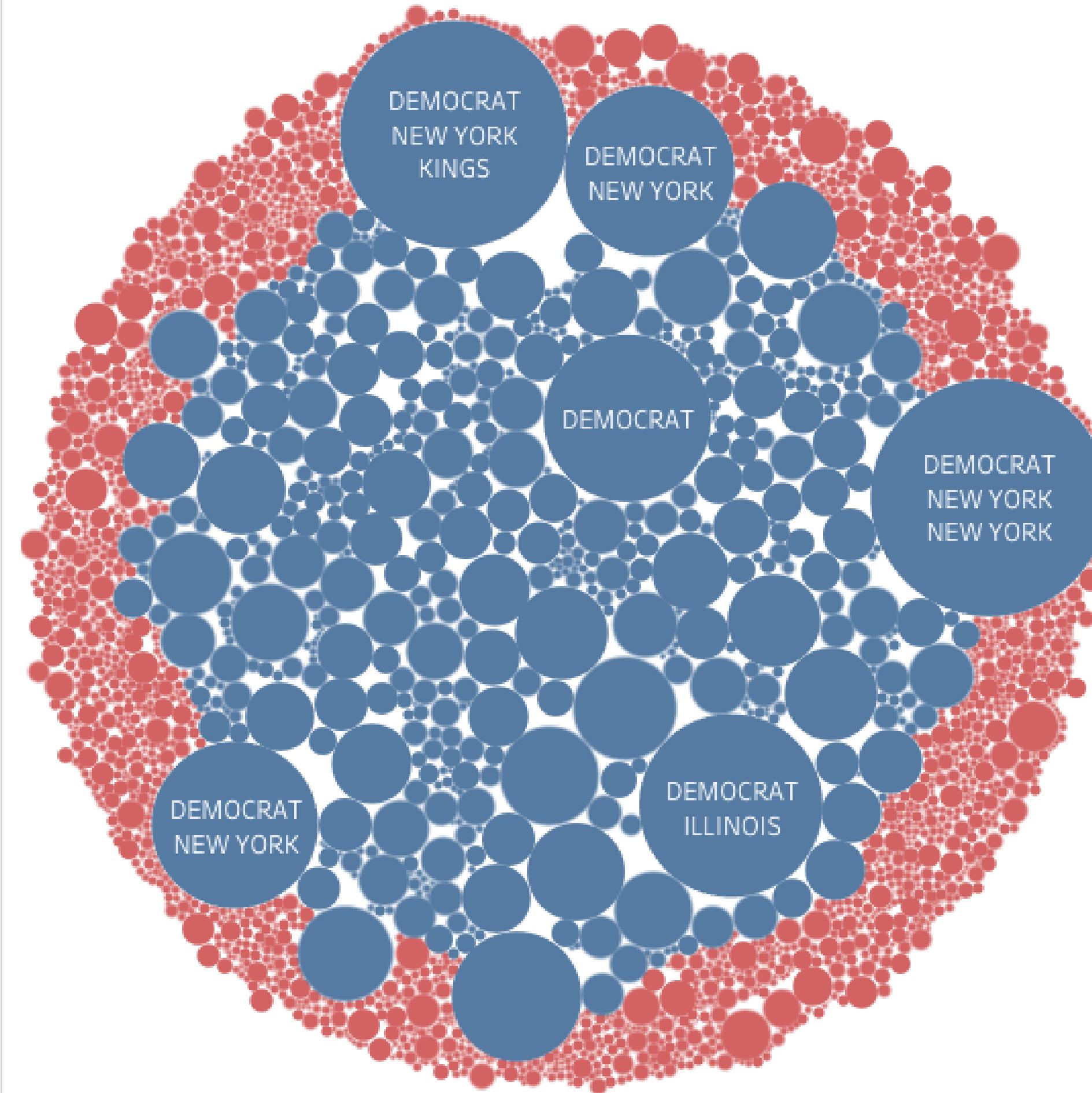
Election outcomes by amount of non-caucasians

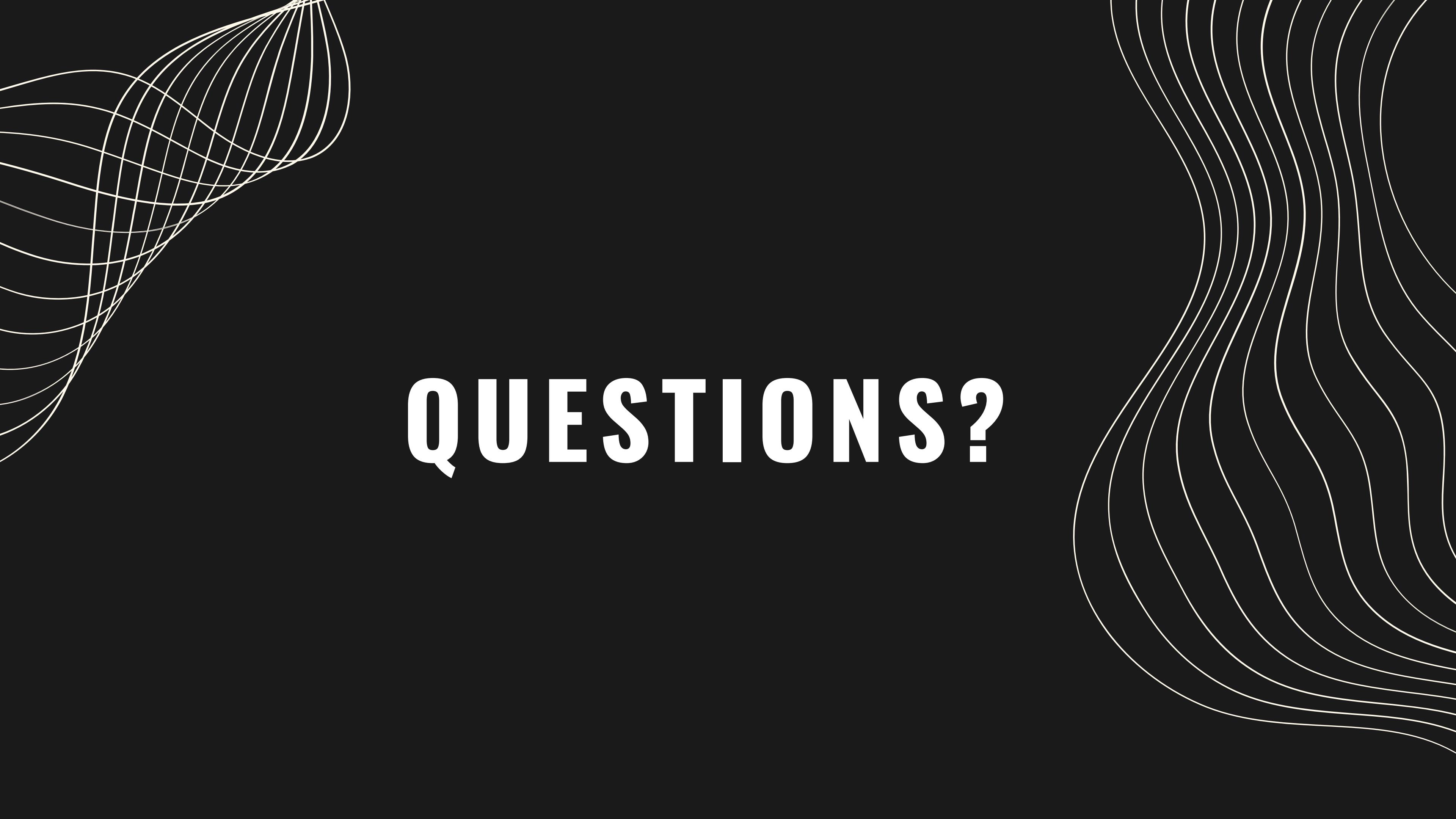


Election outcomes by buildings with <10 housing units



Election outcomes by amount of people with no vehicles





QUESTIONS?



THANK YOU!