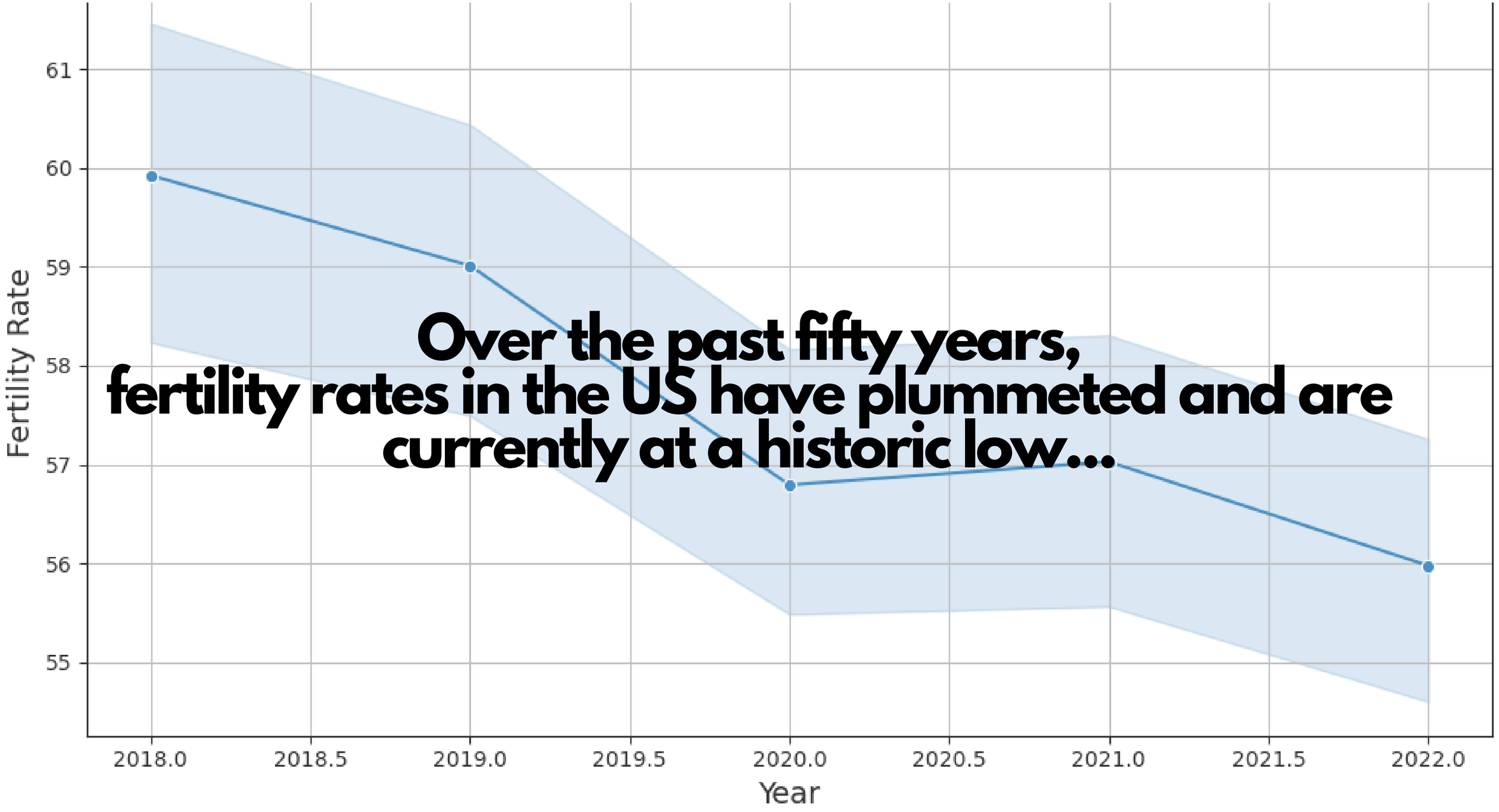
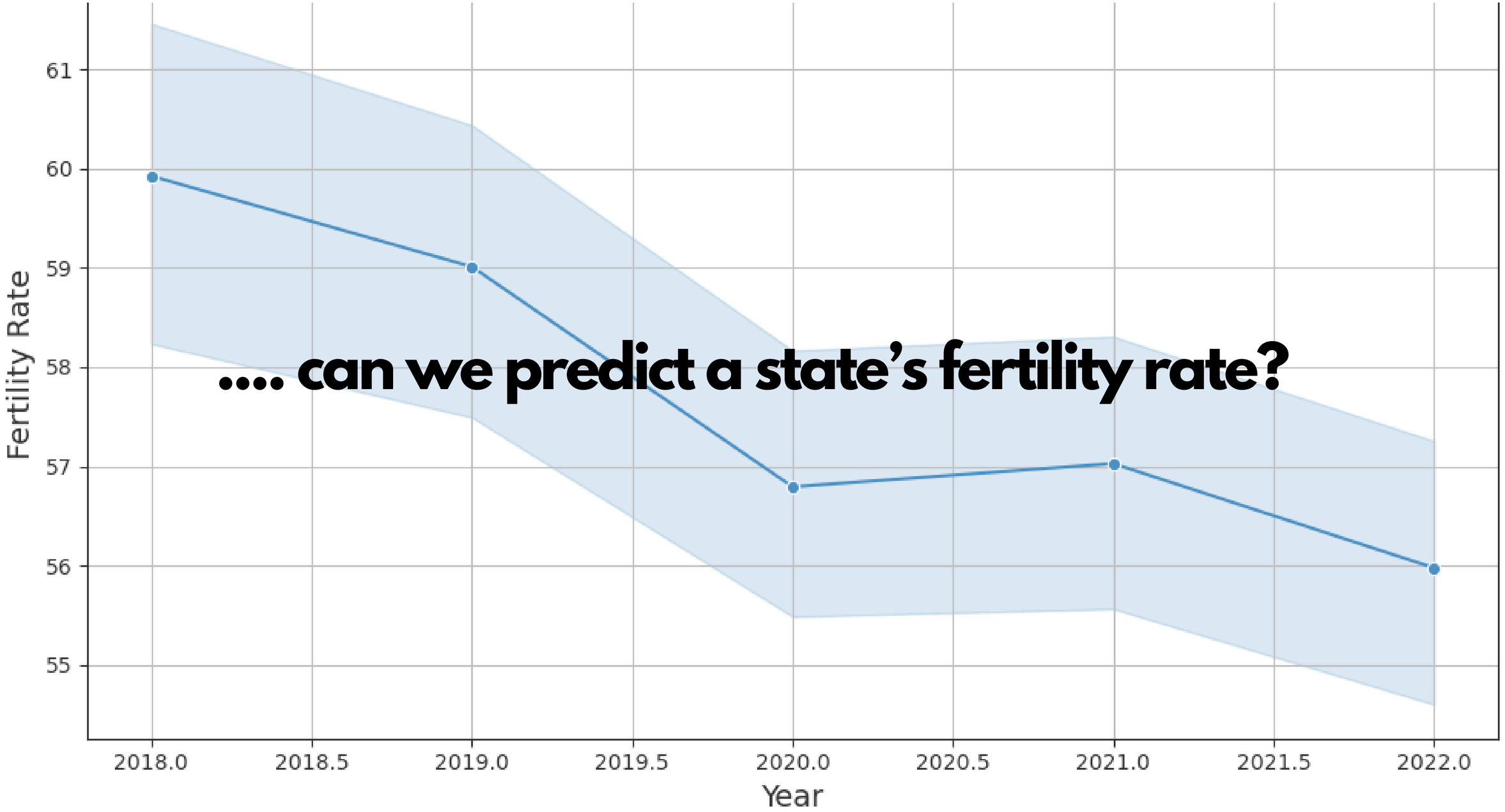
Predicting Fertility Rates in the United States

ISABELLA M., MAIA P., ELIZA G., PETER P.



Over the past fifty years, fertility rates in the US have plummeted and are currently at a historic low...



…. can we predict a state's fertility rate?

We chose to examine 15 predictors across 5 categories for 5 years (2018-2022).

US Census + the ARDA + Cook Political + Law Atlas

[15x250]

RELIGIOUSNESS RATE

COVID YEAR

% WHITE

% BLACK

% HISPANIC

% ASIAN

% MIXED RACE

% FOREIGN BORN

POVERTY MEASURE

% HIGH SCHOOL OR SOME COLLEGE

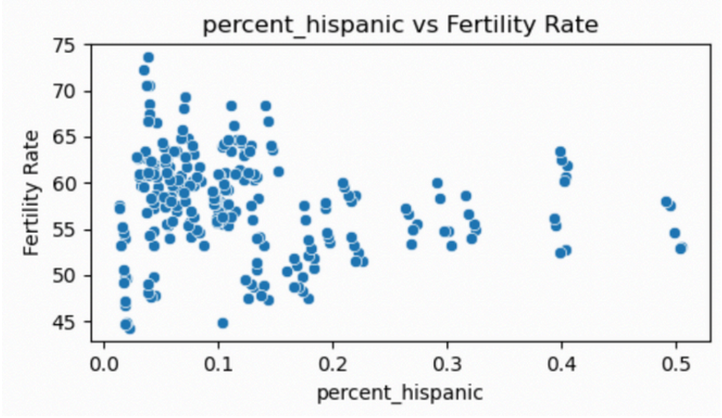
% LESS THAN HIGH SCHOOL

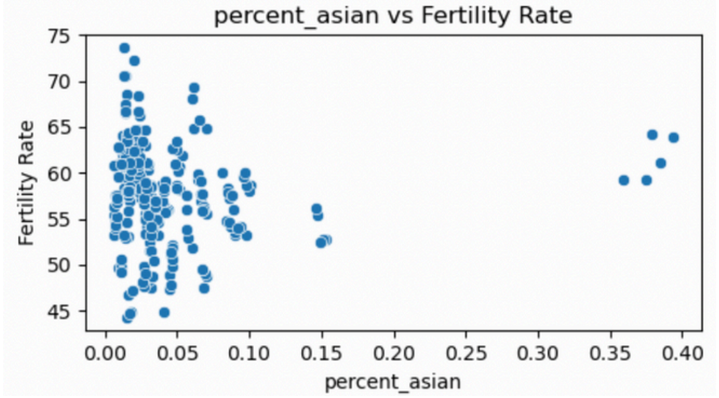
% BACHELORS OR HIGHER

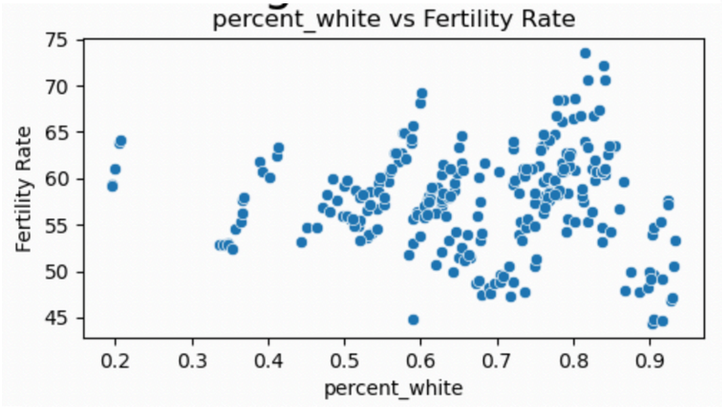
POLITICAL RANKING

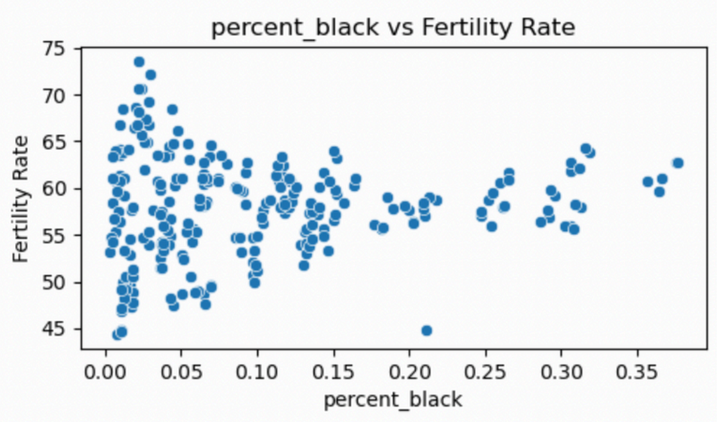
LATES ABORTION

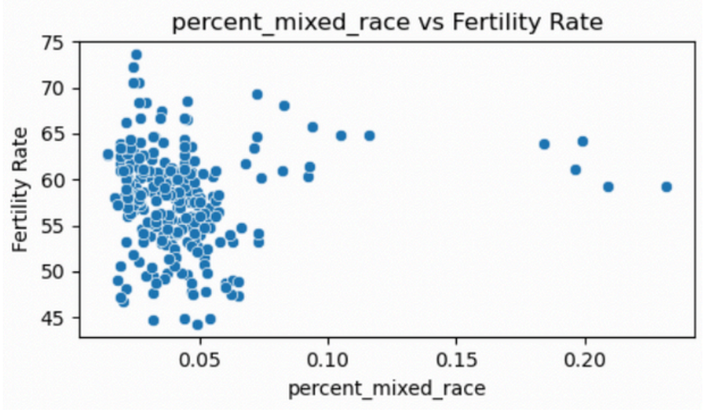
NO ABORTION LAW

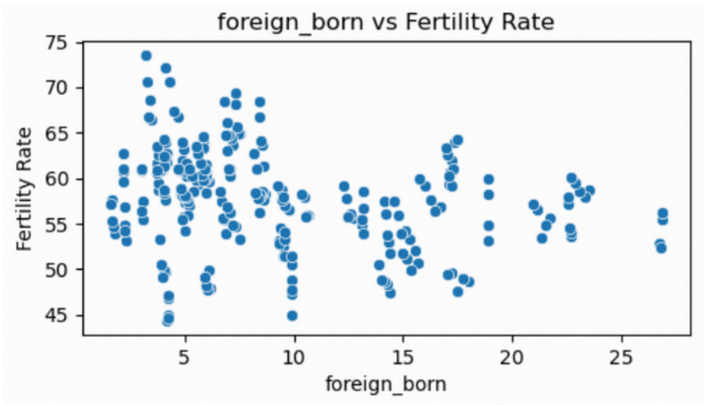




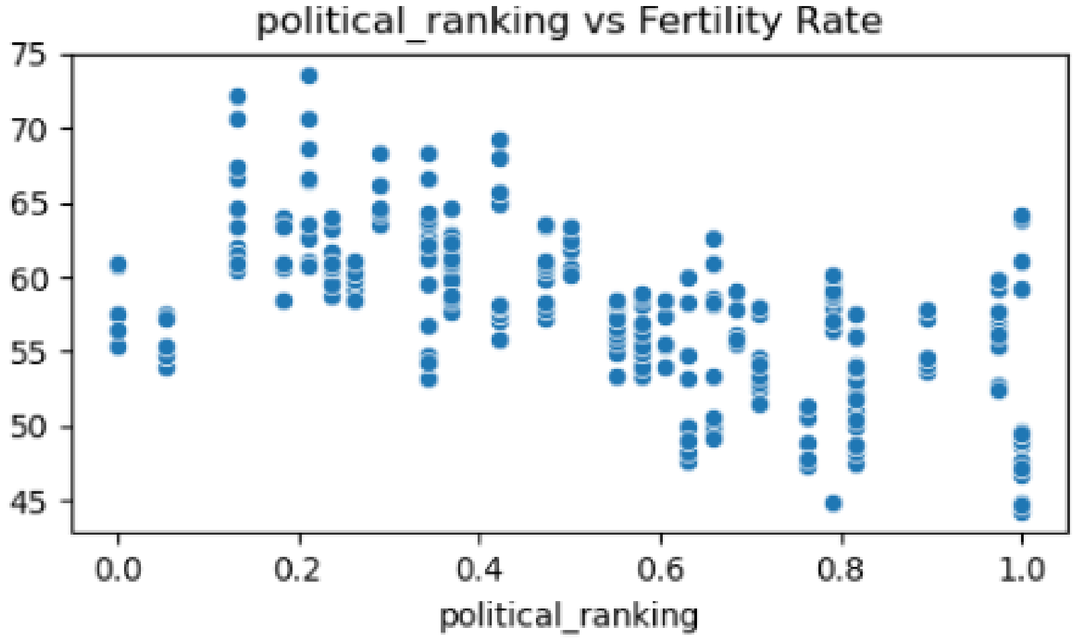


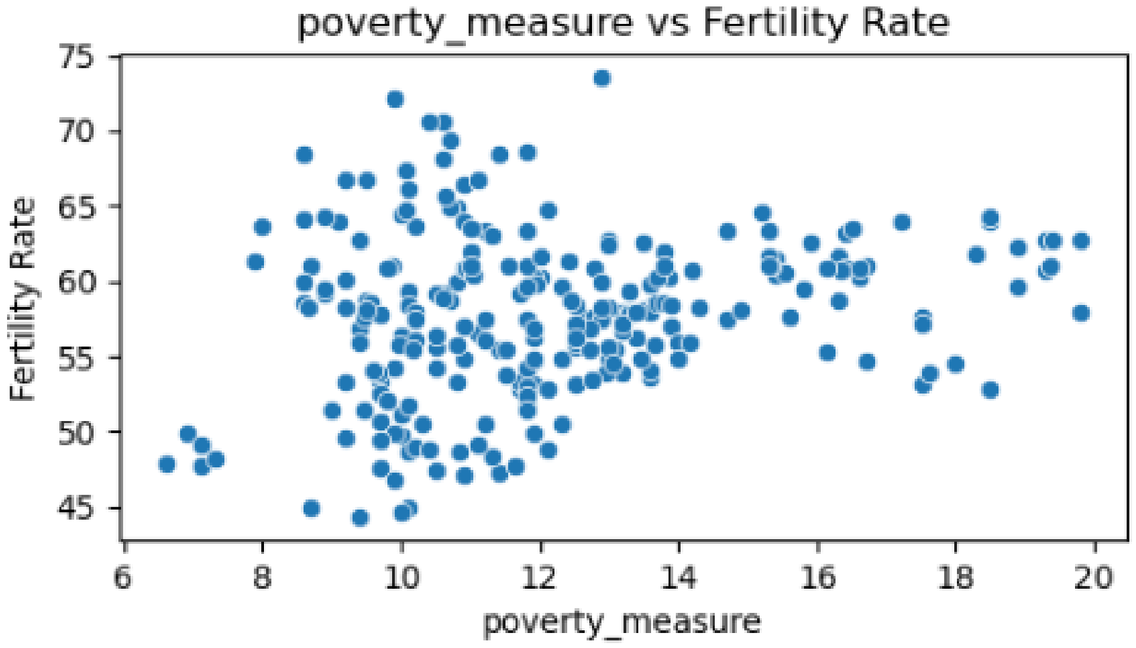






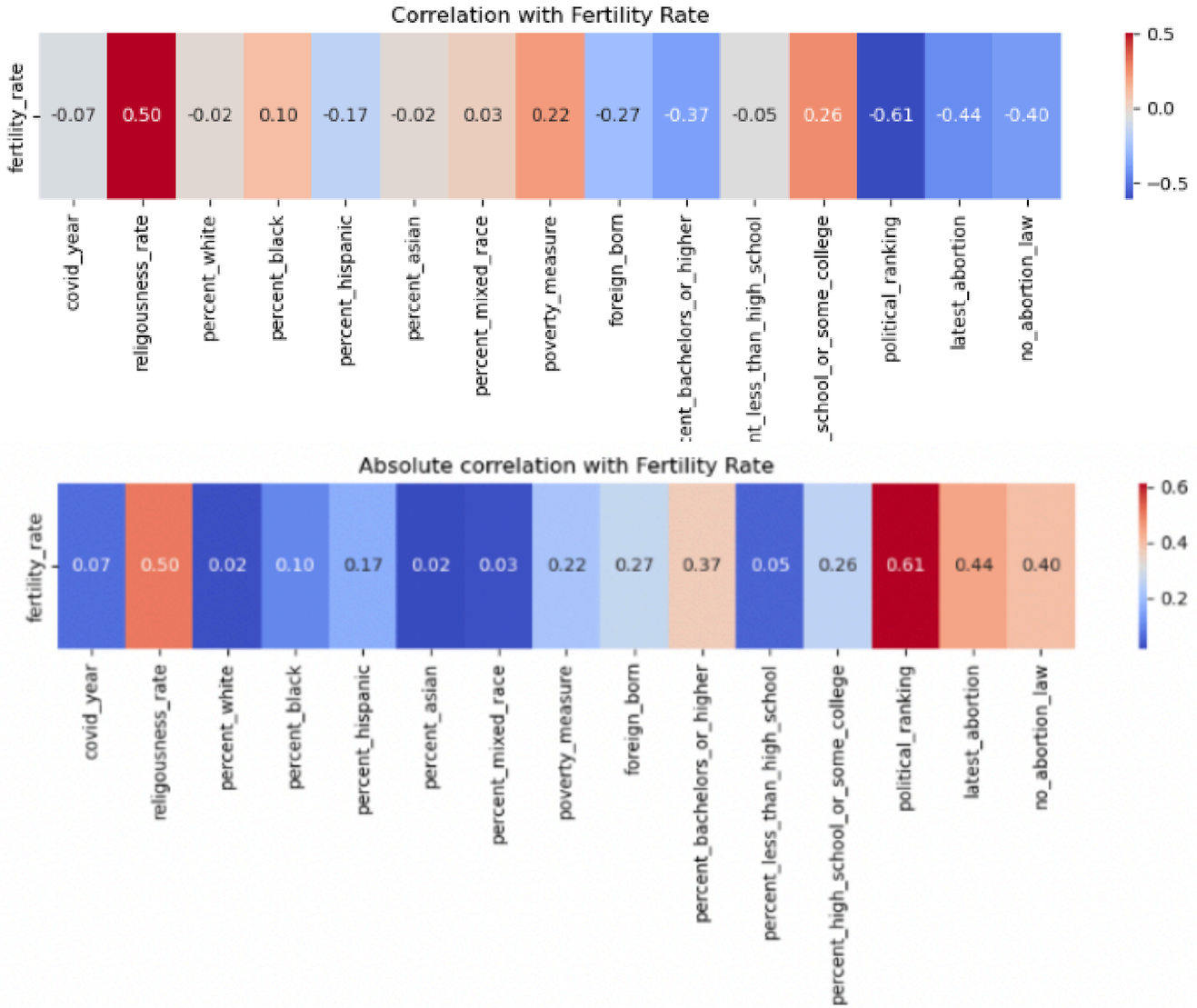
Diversity-related variables show clustering patterns influenced by contextual factors.





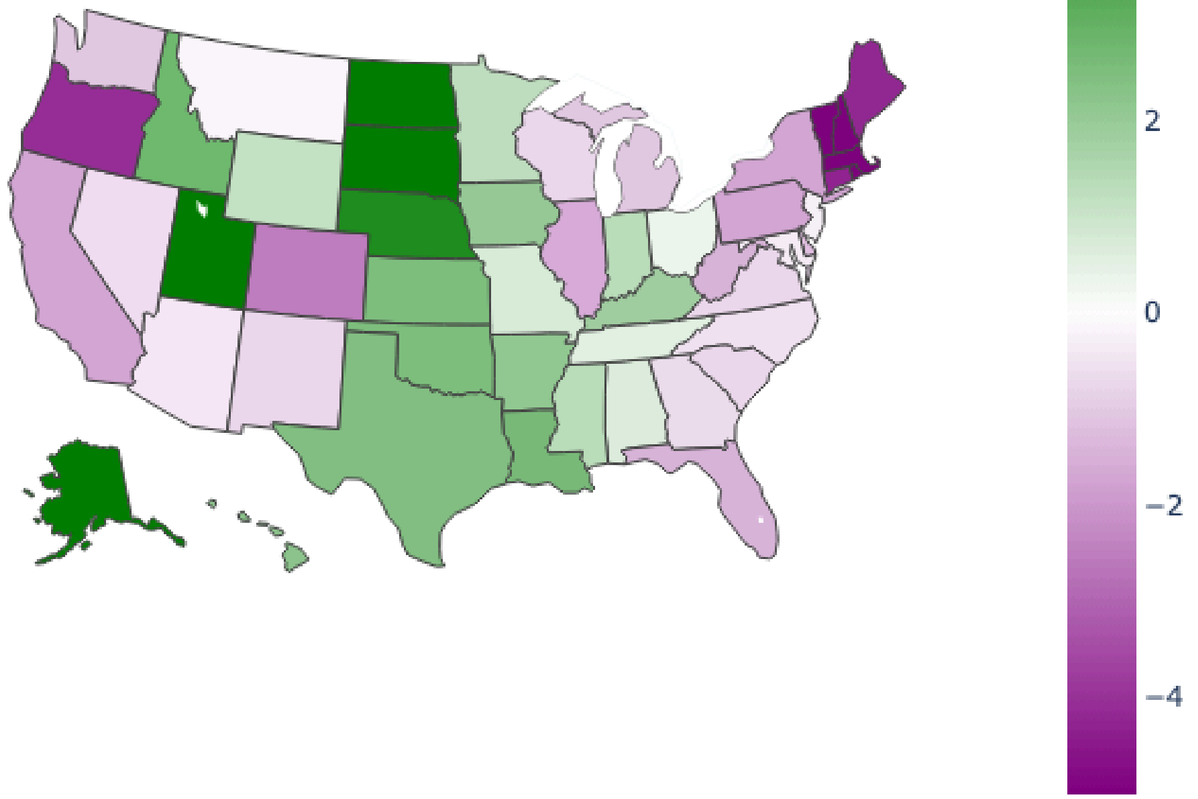
Certain relationships, such as political ranking vs. fertility rate, exhibit linear patterns.

Which predictors correlate with fertility rate most strongly?



How does the fertility rate vary across states?

(94% variance)



Interaction effects likely play a significant role in determining fertility rates.

Cross Validation on 5 Folds

Model 1: Linear Regression with Interaction Terms

CV Mean MSE: 0.946 **R2 CV Mean: -0.149**

0.1 second

we only care about certain predictors

We decided to use lasso and cross validation to identify specifc predictors.

Cross Validation on 5 Folds and StandardScaler

Model 2: Lasso Regression with Interaction Terms

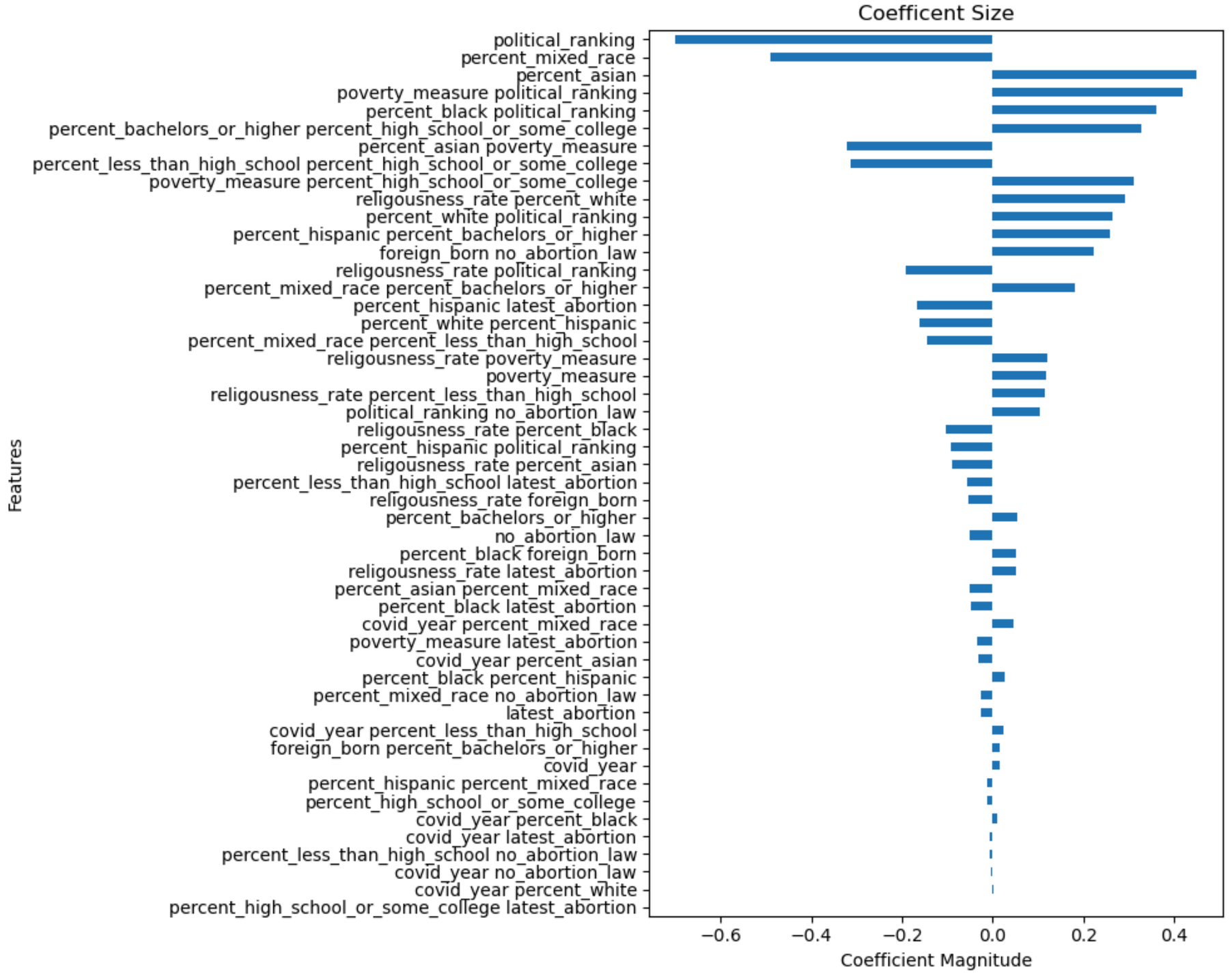
CV Mean MSE: 0.147

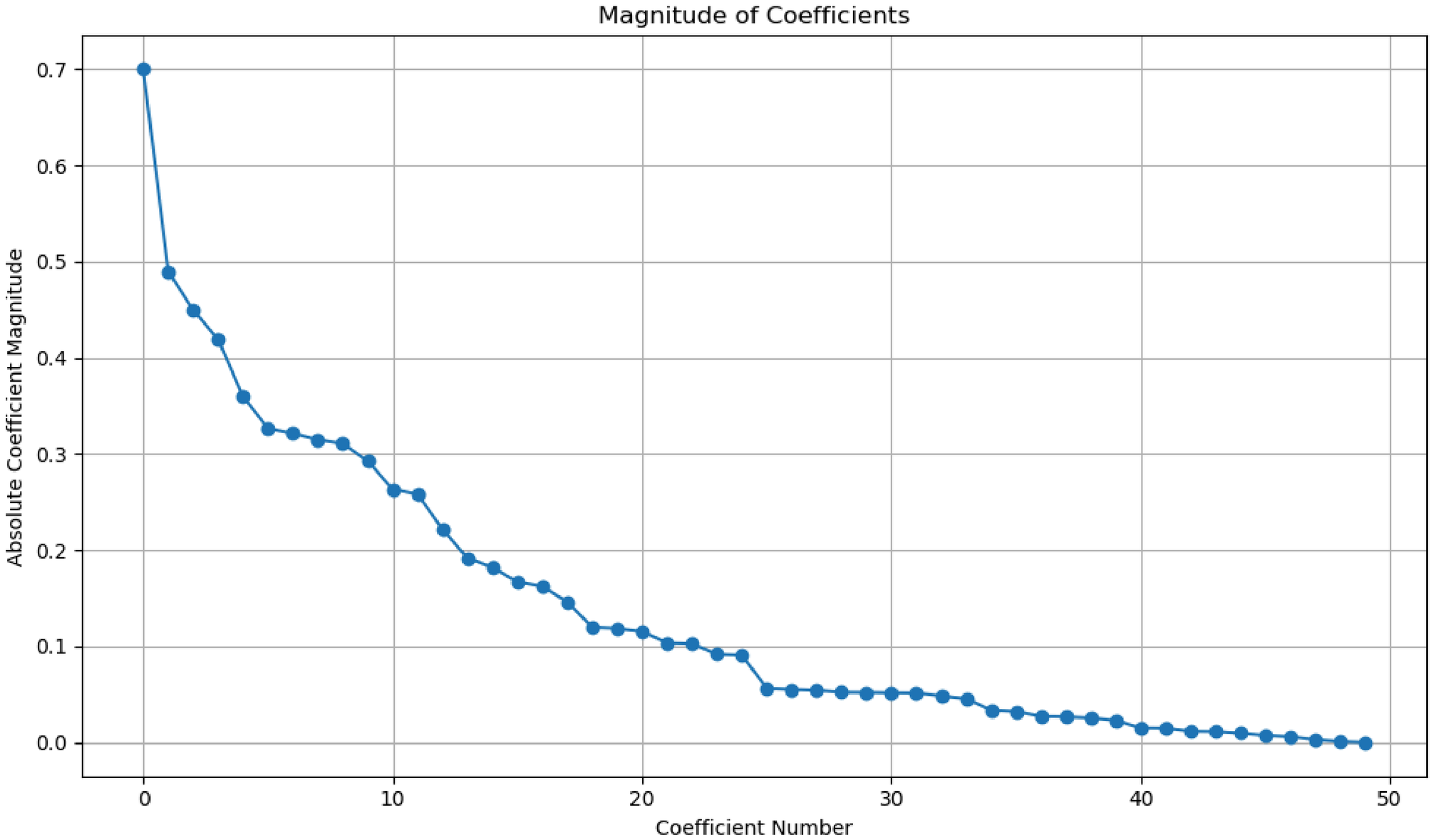
R2 CV Mean: 0.829

2.3 seconds

TEST MSE: 0.115
R2 MSE: 0.888

Both the MSE and R2 values, are similarly good, implying that the model is robust.


Coefficent Size
political_ranking
percent_mixed_race
percent_asian
poverty_measure political_ranking
percent_black political_ranking
percent_bachelors_or_higher percent_high_school_or_some_college
percent_asian poverty_measure
percent_less_than_high_school percent_high_school_or_some_college
poverty_measure percent_high_school_or_some_college
religousness_rate percent_white
percent_white political_ranking
percent_hispanic percent_bachelors_or_higher
foreign_born no_abortion_law
religousness_rate political_ranking
percent_mixed_race percent_bachelors_or_higher
percent_hispanic latest_abortion
percent_white percent_hispanic
percent_mixed_race percent_less_than_high_school
religousness_rate poverty_measure
poverty_measure
religousness_rate percent_less_than_high_school
political_ranking no_abortion_law
religousness_rate percent_black
percent_hispanic political_ranking
religousness_rate percent_asian
percent_less_than_high_school latest_abortion
religousness_rate foreign_born
percent_bachelors_or_higher
no_abortion_law
percent_black foreign_born
religousness_rate latest_abortion
percent_asian percent_mixed_race
percent_black latest_abortion
covid_year percent_mixed_race
poverty_measure latest_abortion
covid_year percent_asian
percent_black percent_hispanic
percent_mixed_race no_abortion_law
latest_abortion
covid_year percent_less_than_high_school
foreign_born percent_bachelors_or_higher
covid_year
percent_hispanic percent_mixed_race
percent_high_school_or_some_college
covid_year percent_black
covid_year latest_abortion
percent_less_than_high_school no_abortion_law
covid_year no_abortion_law
covid_year percent_white
percent_high_school_or_some_college latest_abortion
Features
−0.6
−0.4
−0.2
0.0
0.2
0.4
Coefficient Magnitude


Magnitude of Coefficients
Absolute Coefficient Magnitude
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0
10
20
30
40
50
Coefficient Number

Two of the top five predictors are interaction terms: percent Black x poverty and political leaning x poverty

Of the 50 significant predictors, predictive powers concentrated in the top 24.

Political Ranking

% Mixed Race

% Asian

Poverty Measure x Political Ranking

% Black x Political Ranking

% Bachelors or higher percent high school or some college

% Asian x Poverty Measure

% Less than high school X % Percent high school or some college

Poverty measure X % high school or some college

Religiousness Rate X % White

% White X Political Ranking

% Hispanic X % Bachelors or Higher

Foreign Born X No Abortion Law

Religiousness Rate X Political Ranking

% Mixed Race X % Less than High School

Religiousness Rate X Poverty Measure

Poverty Measure

Religiousness Rate X % Less than High School

Political Rank X No Abortion Law

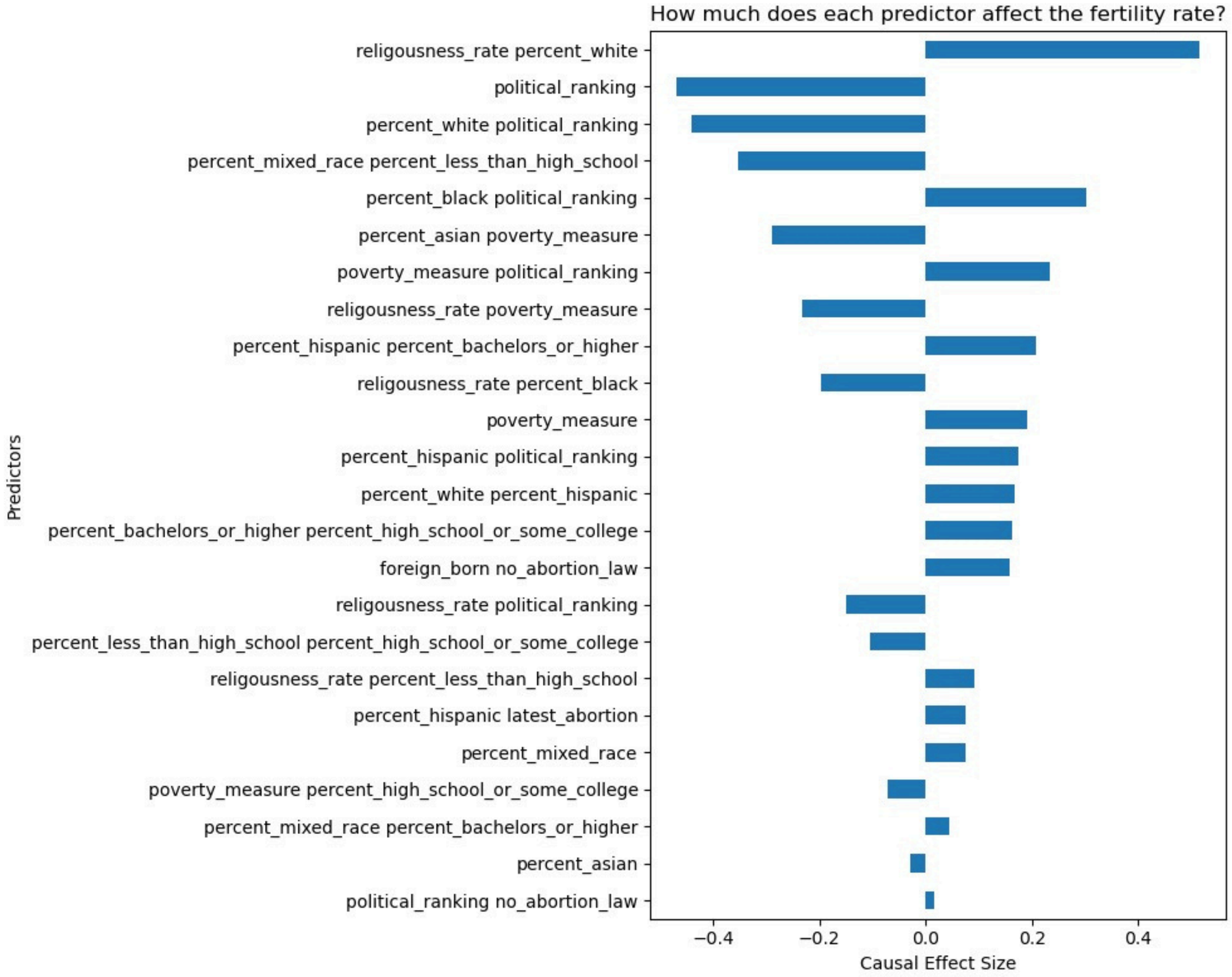
Religiousness Rate X % Black

Focused on top 24 Predictors

Model 3: Causal Inference Analysis

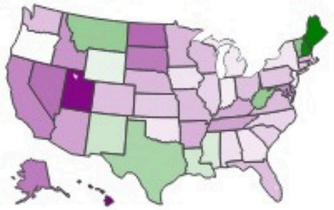
Linear Regression, MinMax scaler,
Propensity Score weighted Ordinary Least Squares

0.0 seconds


How much does each predictor affect the fertility rate?
religousness_rate percent_white
political_ranking
percent_white political_ranking
percent_mixed_race percent_less_than_high_school
percent_black political_ranking
percent_asian poverty_measure
poverty_measure political_ranking
religousness_rate poverty_measure
percent_hispanic percent_bachelors_or_higher
religousness_rate percent_black
poverty_measure
percent_hispanic political_ranking
percent_white percent_hispanic
percent_bachelors_or_higher percent_high_school_or_some_college
foreign_born no_abortion_law
religousness_rate political_ranking
percent_less_than_high_school percent_high_school_or_some_college
religousness_rate percent_less_than_high_school
percent_hispanic latest_abortion
percent_mixed_race
poverty_measure percent_high_school_or_some_college
percent_mixed_race percent_bachelors_or_higher
percent_asian
political_ranking no_abortion_law
Predictors
−0.4
−0.2
0.0
0.2
0.4
Causal Effect Size

Presence of these effects in the United States

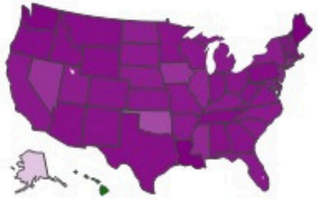
religousness_rate percent_white



political_ranking

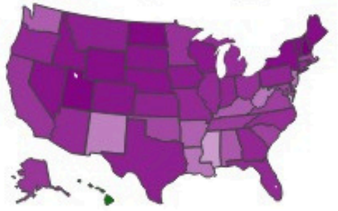
percent_white political_ranking

percent_mixed_race percent_less_than_high_school

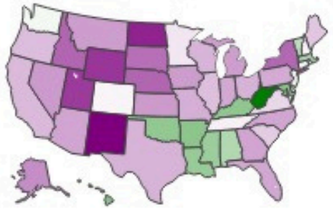


percent_black political_ranking

percent_asian poverty_measure

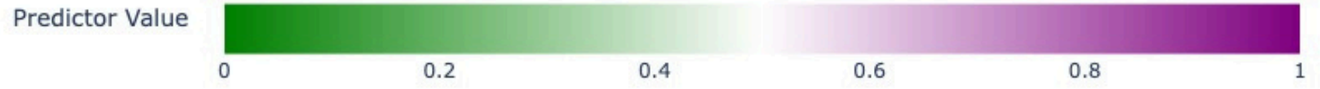


poverty_measure political_ranking

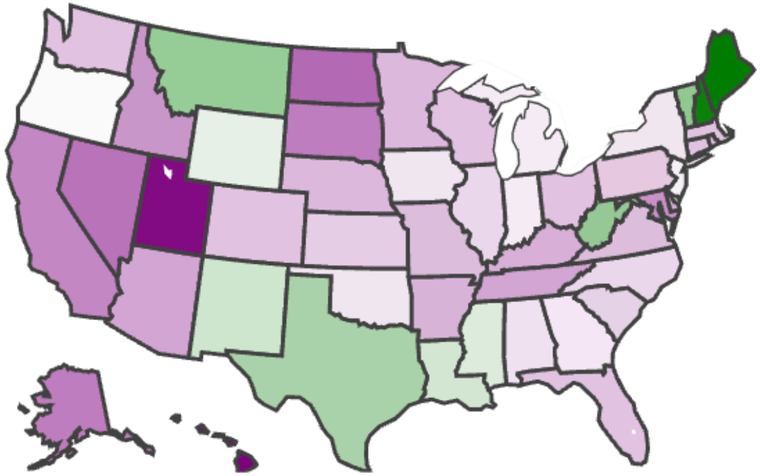


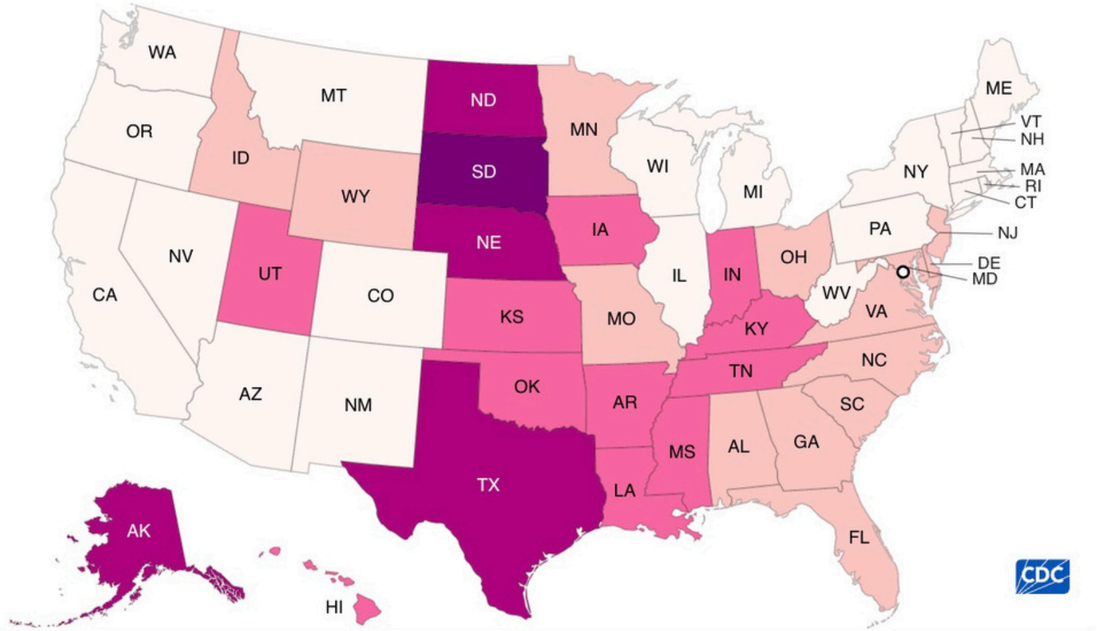
religousness_rate poverty_measure

percent_hispanic percent_bachelors_or_higher



religousness_rate percent_white





Fertility Rate[1]

- 44.3 - 54.9
- 55.3 - 59.1
- 59.2 - 61.8
- 61.8 - 65.3
- 65.4 - 92.8

AREAS FOR IMPROVEMENT

While examining things on the state level was useful, we would use additional time to look at the county or zip-code level.

Our dataset is limited in size, both in number of observations (rows) and number of predictors (columns).

This scarcity of predictors may mean we are missing out on significant confounding variables, important for causal inference.