

# Hate Crimes by State

*Theodore Dounias*

*November 13, 2017*

## Motivating Questions

What are the key variables that we can use to predict the number of hatecrimes in a certain state?

What is the best predictor available?

Does voting for Donald Trump correlate with increased hatecrimes in a certain state?

What predictors fail/succeed contrary to popular belief?

## Purpose

Lay the foundation for further study of certain predictors, create policy solutions to affect change.

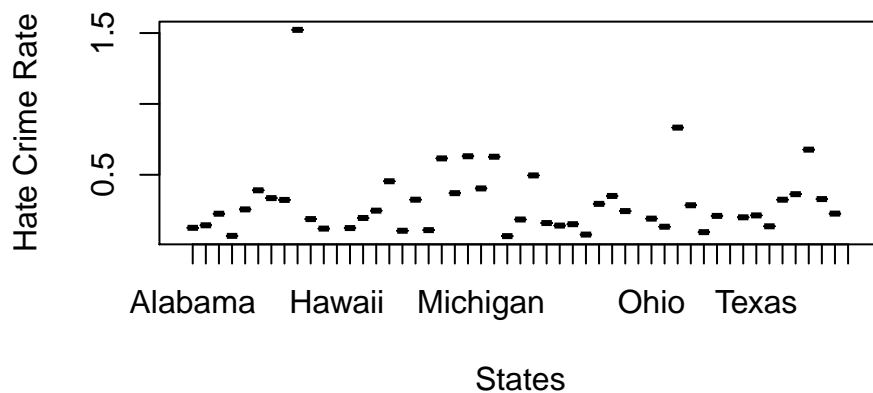
## Data Source

The data we will be using is a dataset from the fivethirtyeight GitHub page. This data has been tidied and wrangled already. To this we might wish to add census data for each state to check on the predictive value of other variables—especially binary ones concerning status of policy on minority issues that are lacking in this dataset.

## Data Structure

The unit of this data is the individual state. For each state, the number of hatecrimes per 100k residents is recorded, along with variables including race, income, vote, and education. Census data, if included, would also be structured by state, and would have to be joined with the initial dataset.

```
plot(d[, 1], d[, 11], xlab = "States", ylab = "Hate Crime Rate")
```



## Model Proposed

Our goal here is prediction and predictive accuracy. To accomplish this, we will fit a variety of models to find the most accurate one available. This will include a tree-based algorithm and multiple linear regression with the use of LASSO and Ridge methods.