approximately normal

Predicting 2018 Voter Turnout Rates in Minnesota

Agenda

Data

EDA

Methodology

Prediction

Fun Facts

Our Data

Data Sources and Features

Minnesota Secretary of State

 ⇢ General election results
   - 2010, 2012, 2014, 2016

| Variable Name	Variable Description
pctcode	Precinct Number (unique within county)
countycode	County Number
congdist	US Representative (Congressional) District
reg7am	Number of pre-registered voters in precinct
edr	Number of voters registering when they vote
…	…

Federal Reserve Economic Data (FRED)

 ⇢ Obtained using the FRED API
   - 2010 - 2017
   - monthly/ yearly
   - county level

Variable Name	Variable Description
income_med	Estimate of Median Household Income
homeowner	Homeownership Rate
inc_ineq	Income Inequality
burd_house	Burdened Households
age_med	Median Age of the Population
…	…

EDA

Variation across Precincts in Number of Voters
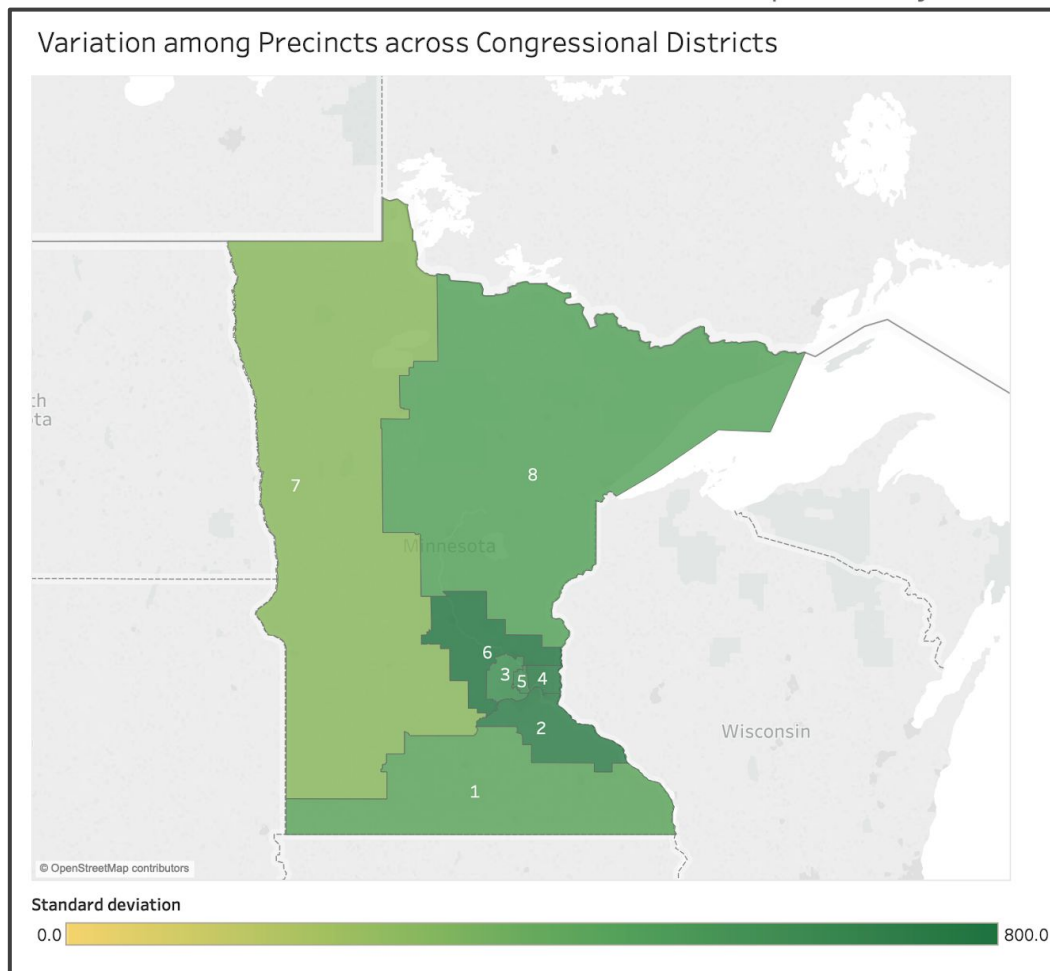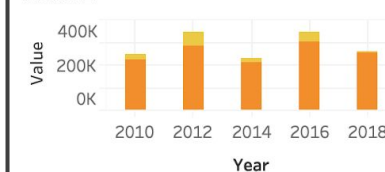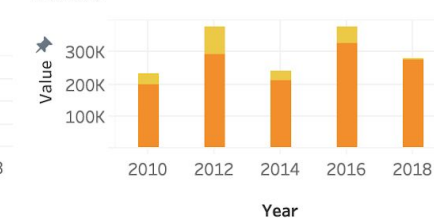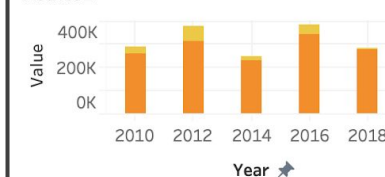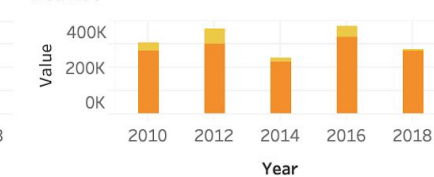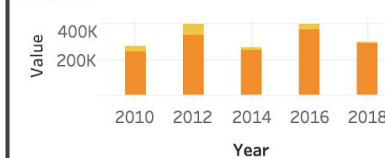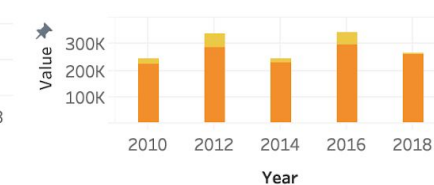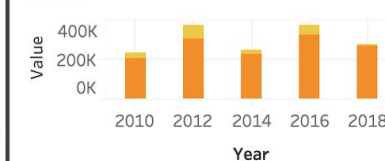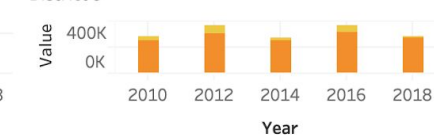
Change in Total Number of Voters Over Time

Characteristics Likely to Contribute to the Outcome

Top Factors Supporting Democratic Party
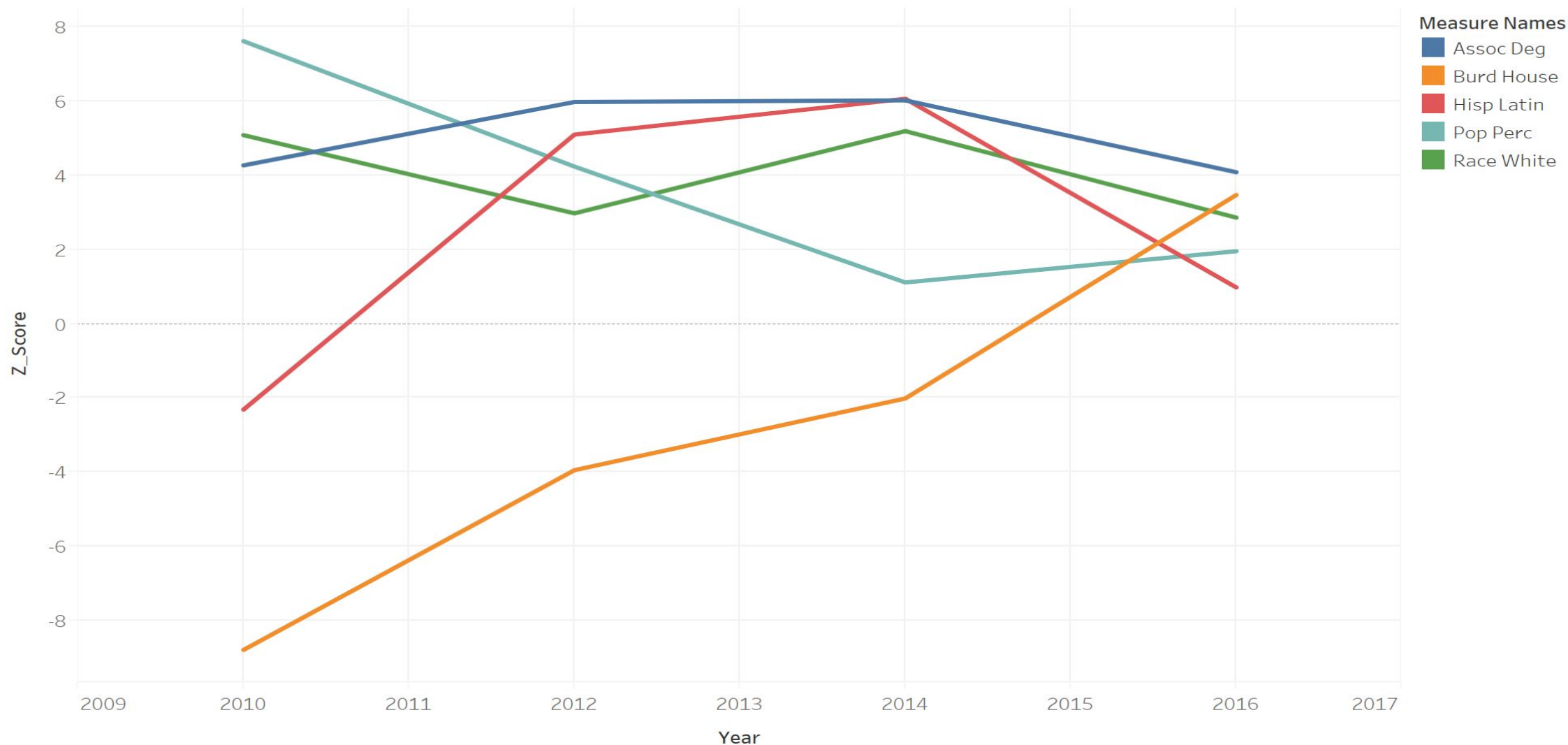
Characteristics Likely to Contribute to the Outcome

Top Factors Supporting Republican Party
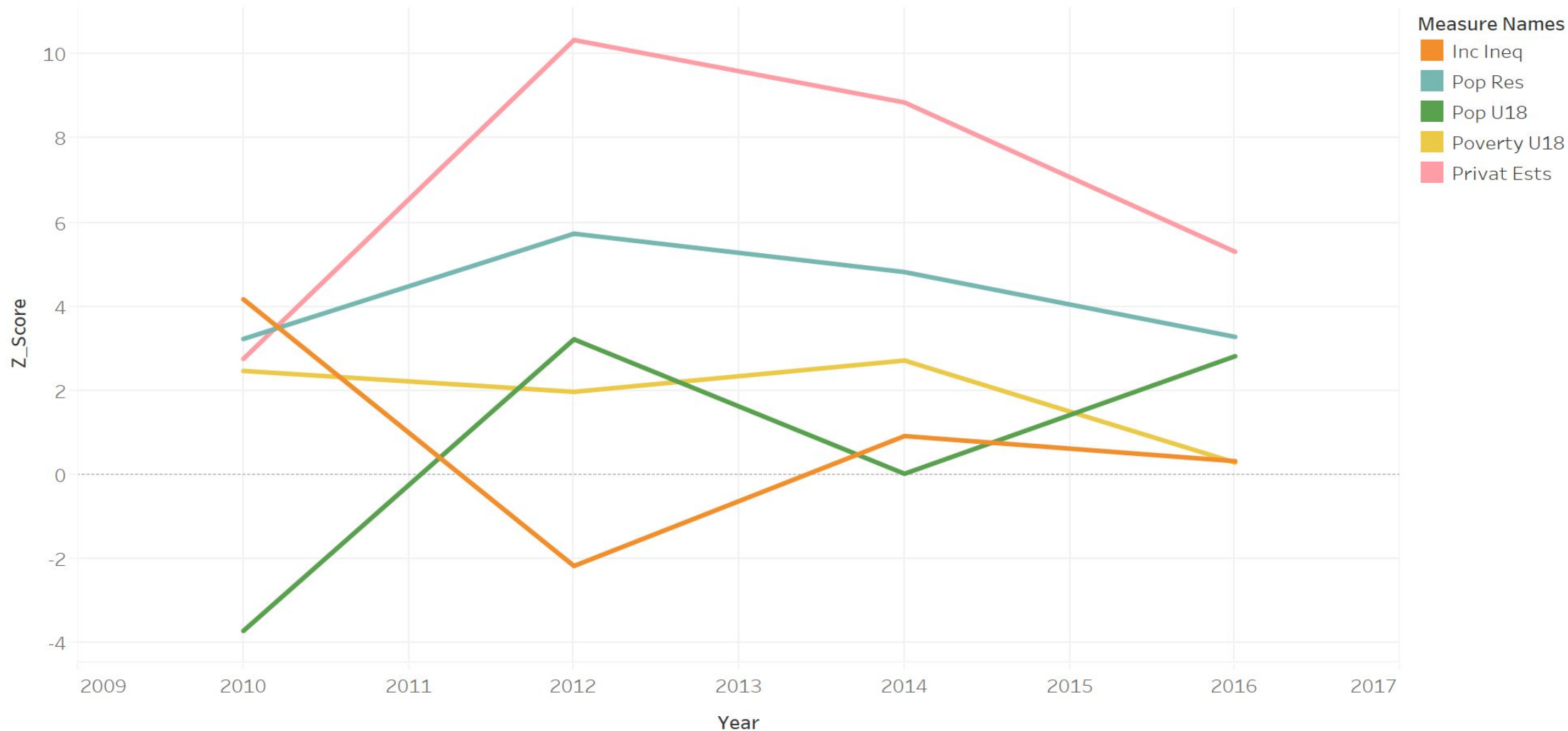
Our Model

Formula

$$V_i = \sum_{j \in i} edr_j + (\Delta_j * reg7am_j)$$

Where:

- i represents each of the 8 congressional districts
- j represents each of the precincts that belongs to the congressional district
- edr represents the number of same-day registrations
- reg7am represents the number of registered voters as of 7:00 AM on election day
- Δ represents the percentage of pre-registered voters that vote

Regression Tree

Random Forest

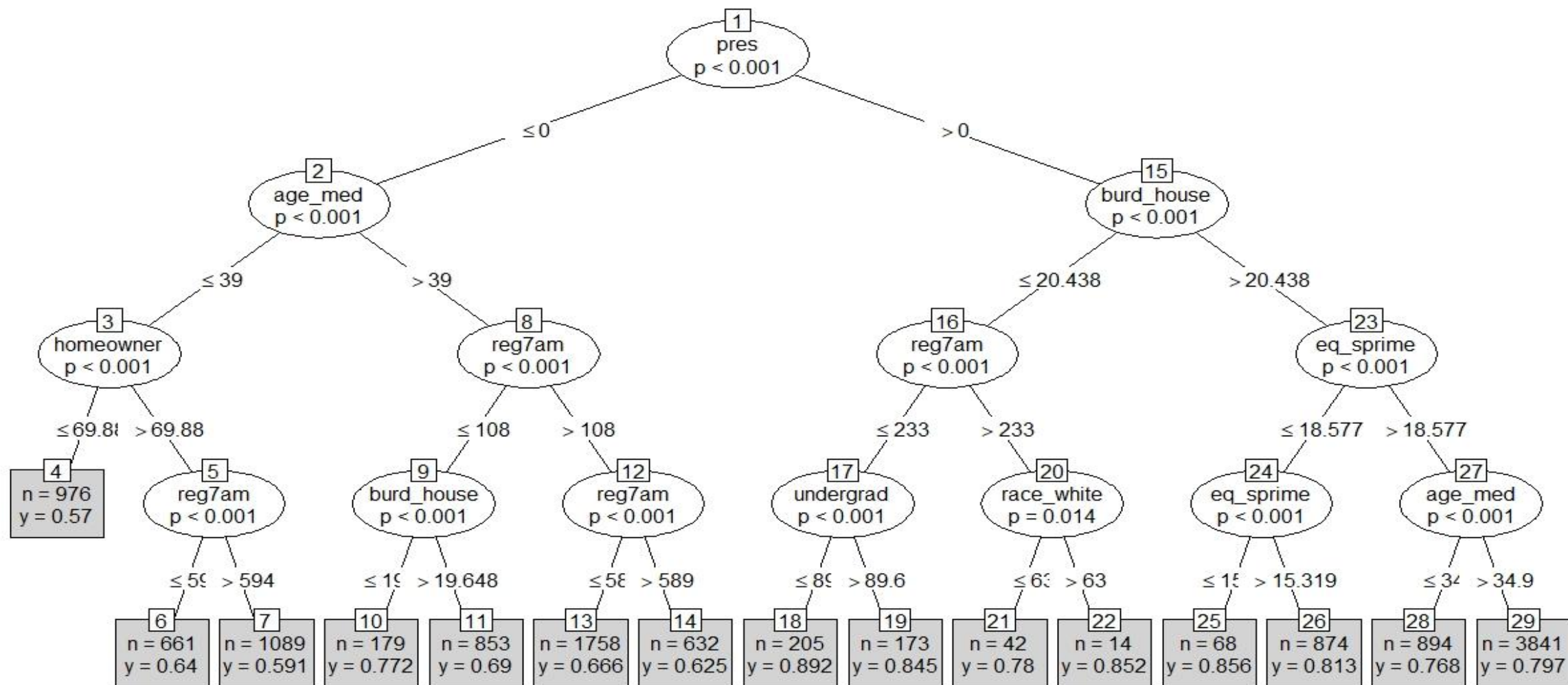**rforest_mod**
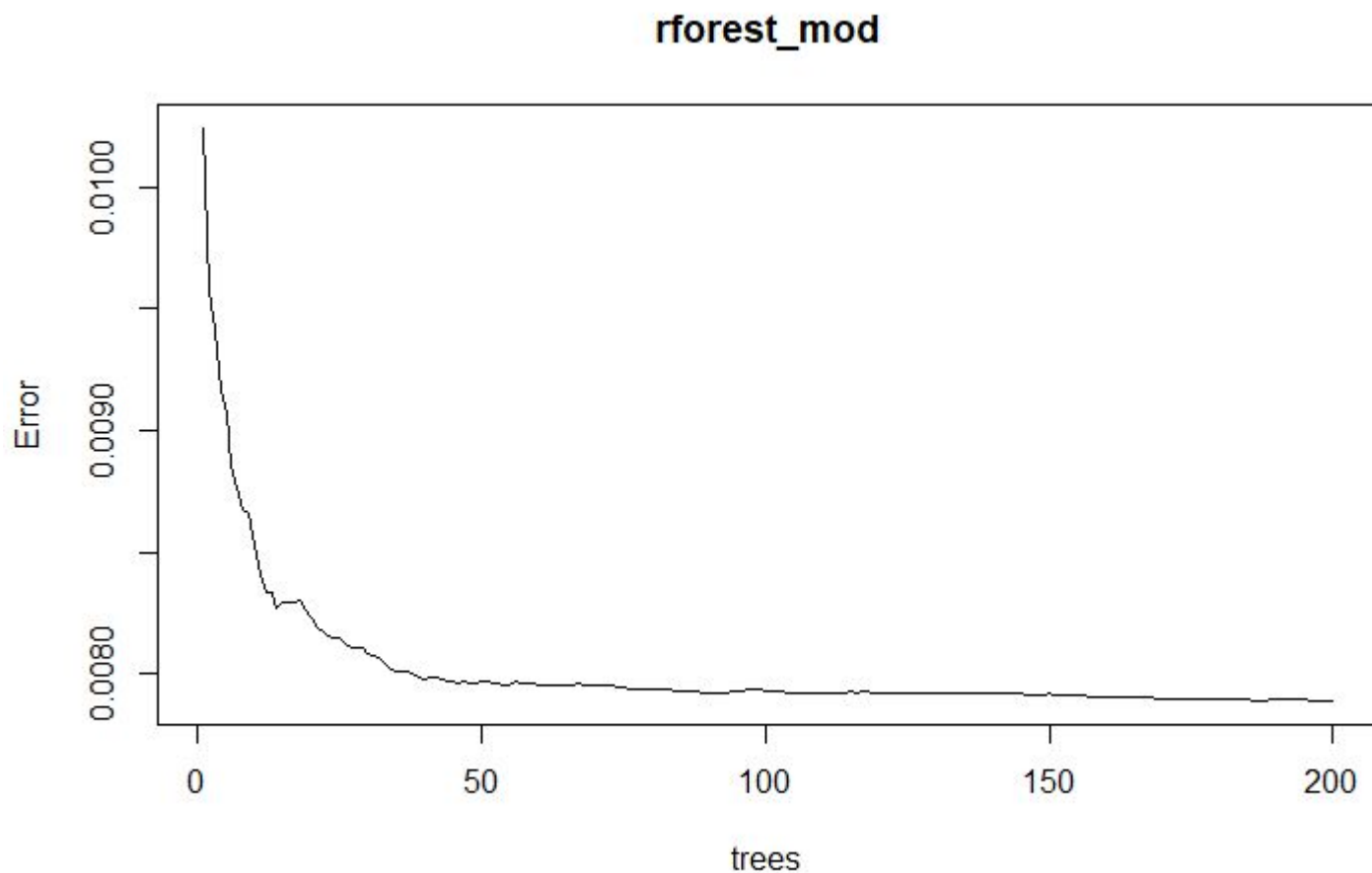
→ Trained using R package `randomForest`

- Test set: 2014 and 2016 precincts
- 10 features sampled at each node
- 200 trees generated
- Predicted both edr and Δ

→ Improved Accuracy

- Reduced error over linear model by 30%
- MAE on test set: 0.0603

Factors that Appear to Influence Voter Turnout

Feature Relation to Turnout
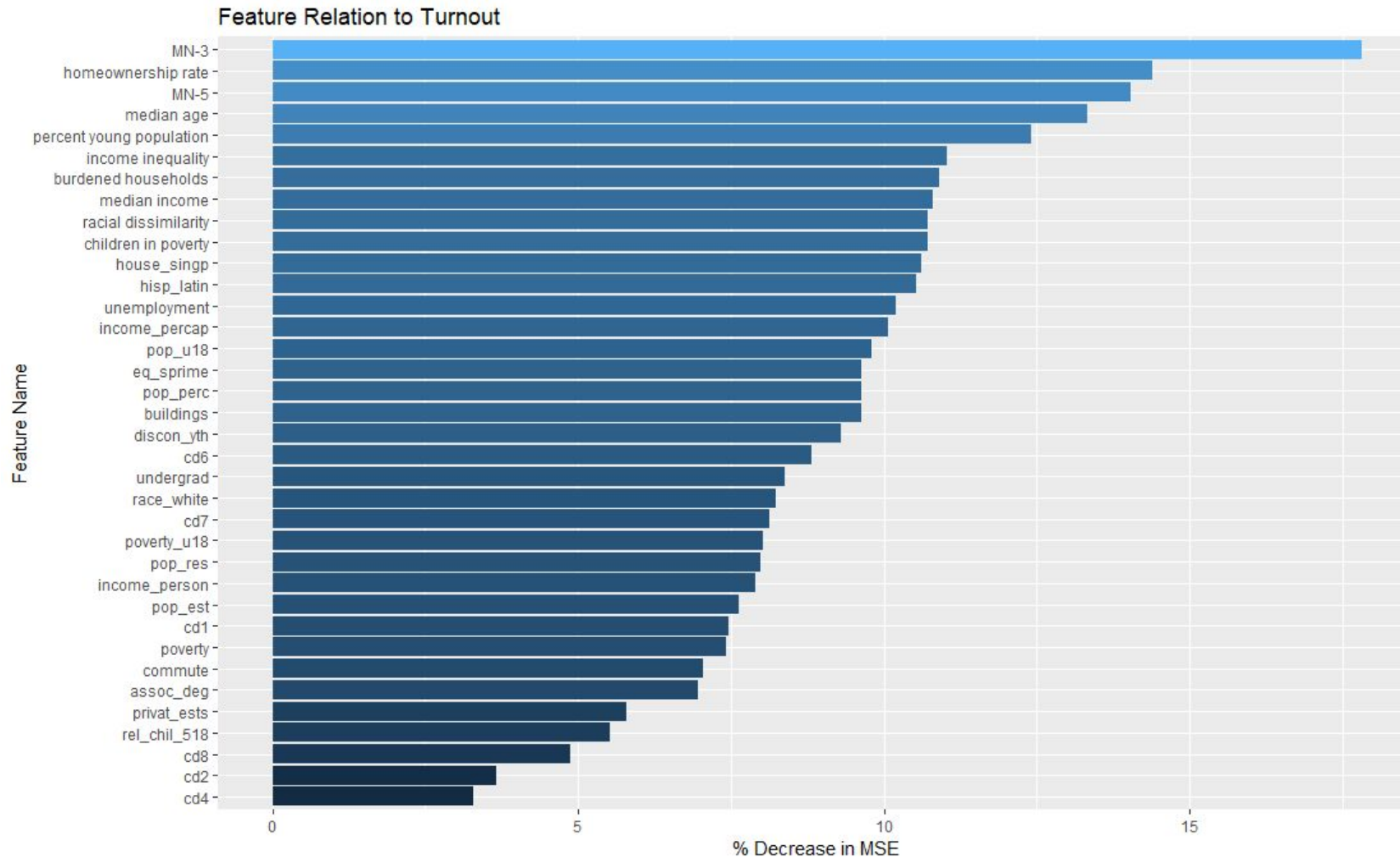
Top Five Factors

→ 3rd Congressional District

→ Home Ownership Rate

→ Income Inequality

→ Median Age

→ Racial Dissimilarity

Our Prediction

MN-1	258771
MN-2	282804
MN-3	305077
MN-4	275145
MN-5	279661
MN-6	276054
MN-7	263741
MN-8	281784

Predicting Data (2018)

Hennepin County Unemployment Rate
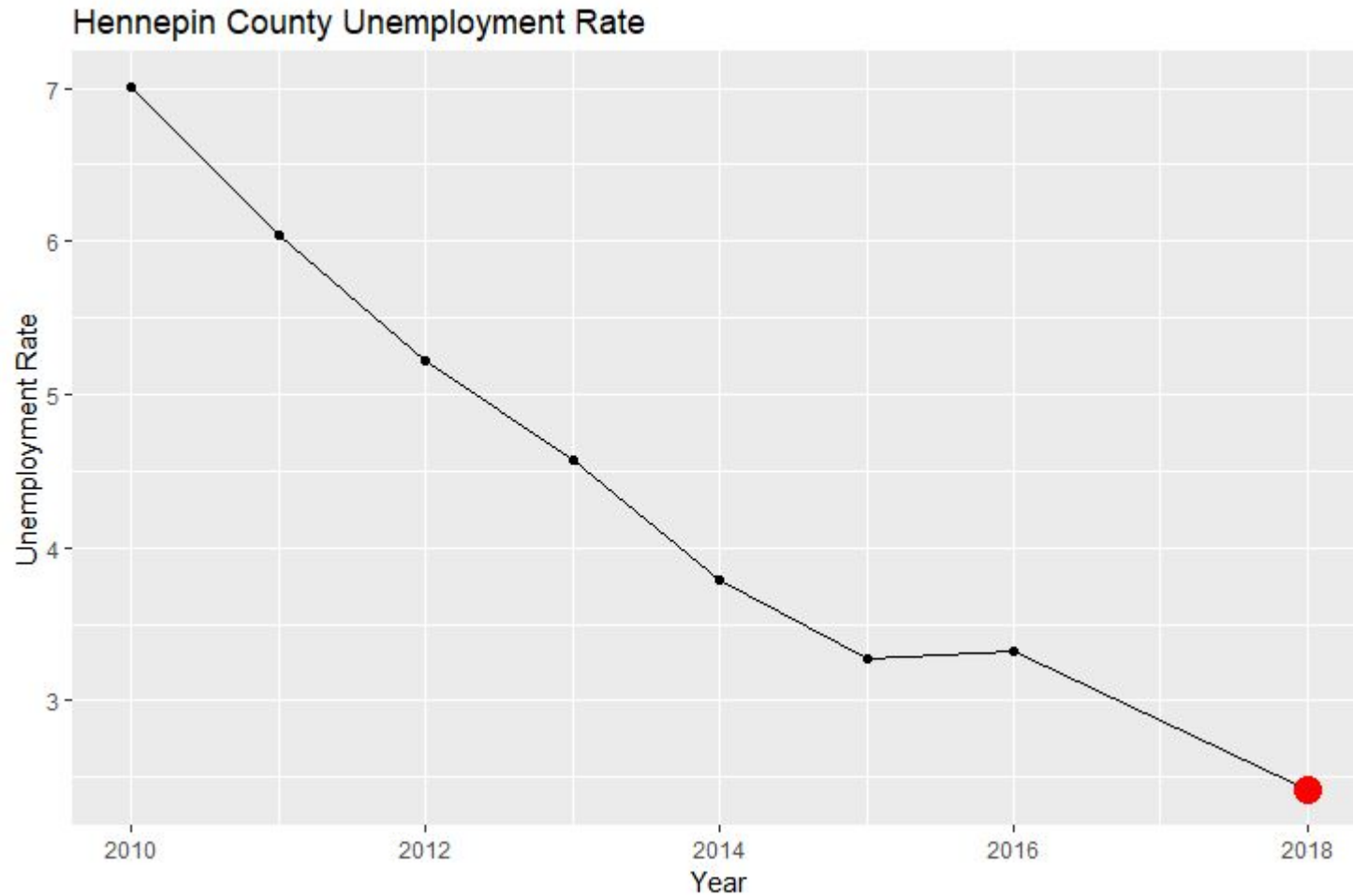
→ Estimated based on historical trends

- Weighted average of bi-annual difference
- Higher weights assigned to recent years

Predictions

Total Voting Ballot Prediction on Congressional District Levels
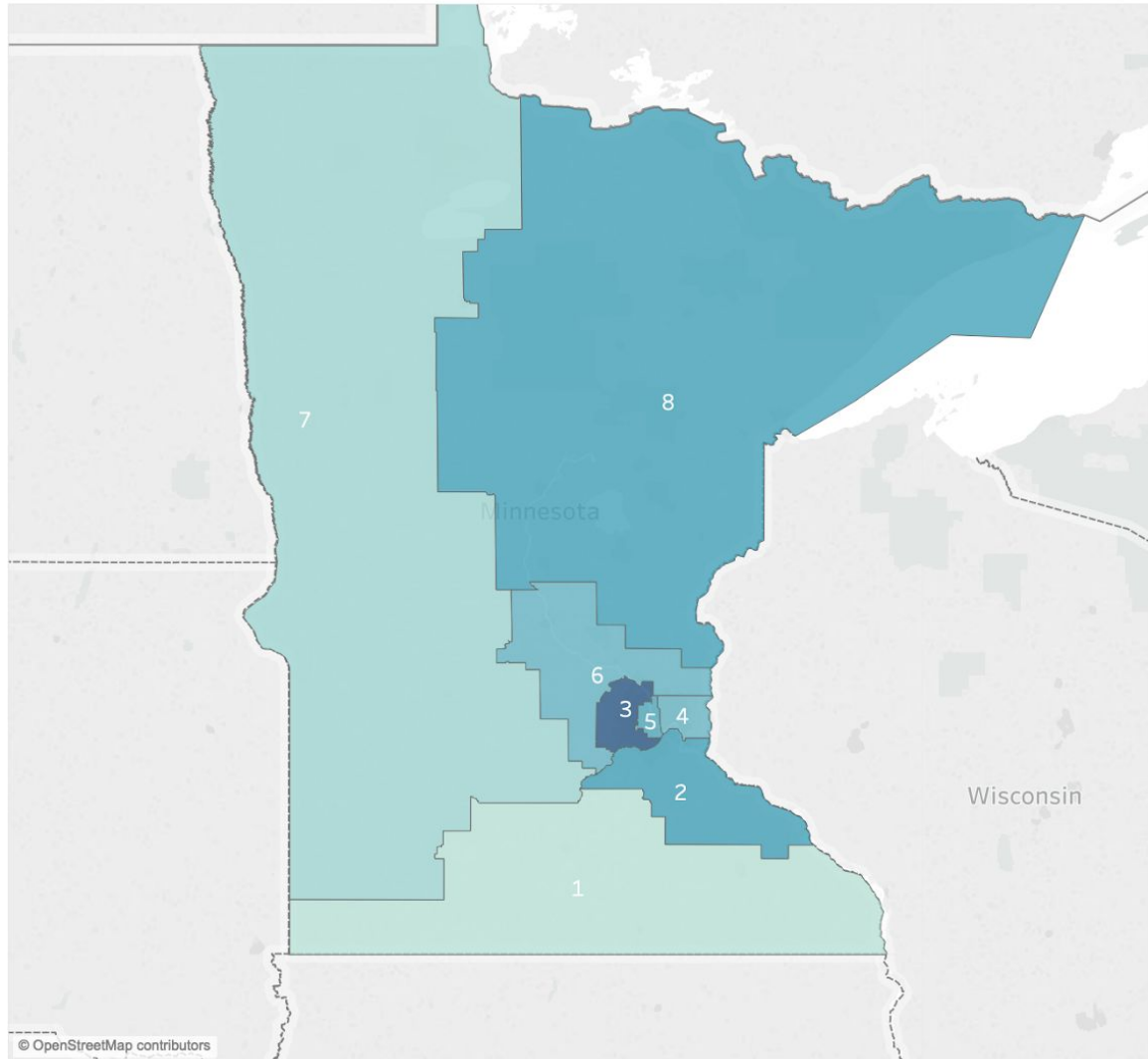

Congvotes

258771 ▭ 305077

Turnout rate Prediction on Congressional District Levels

Turnout

0.66000 ▭ 0.69526

© OpenStreetMap contributors
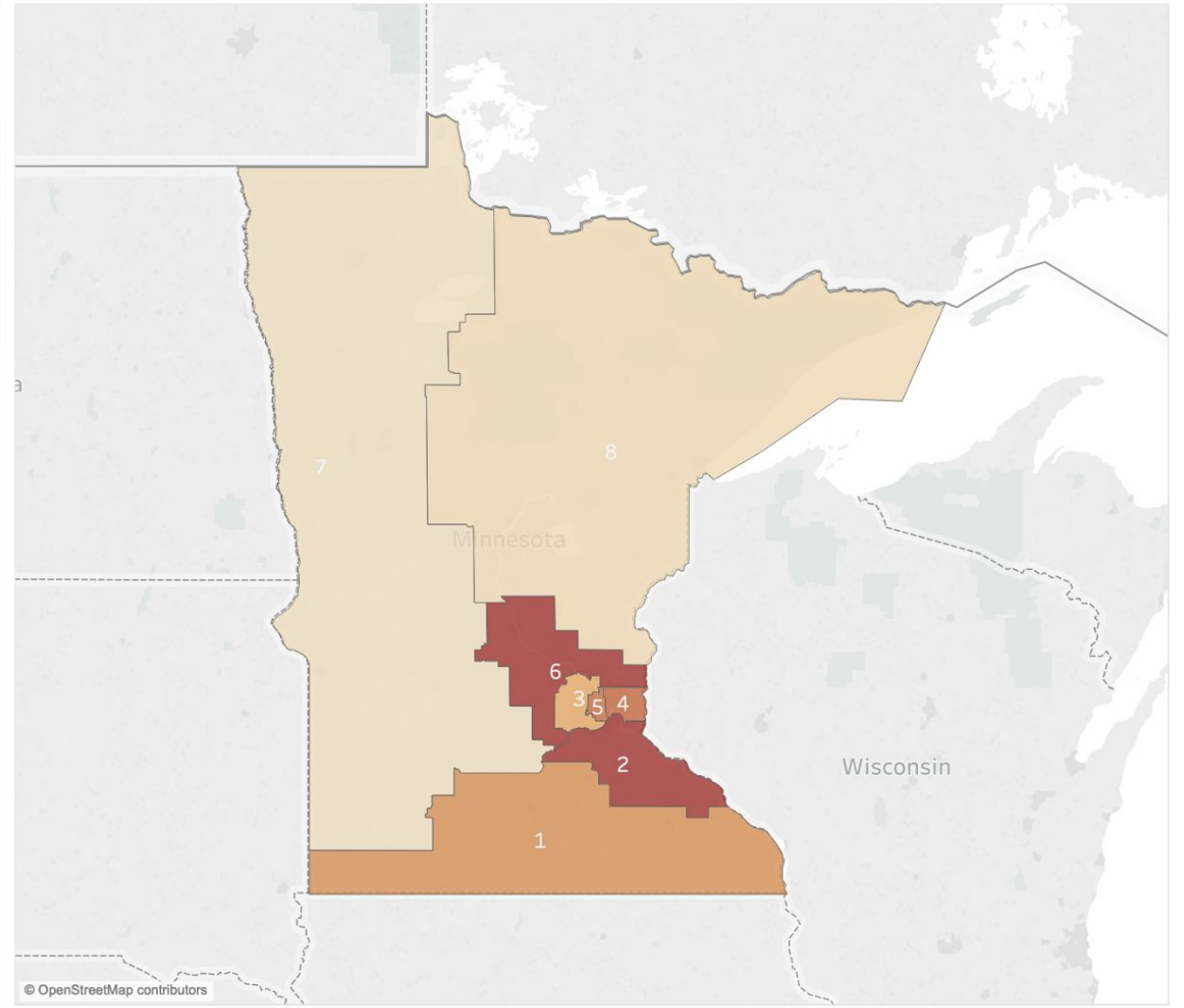
Interesting Findings

Correlation between YouTube Views & Voter Turnout

YouTube Views

⇢ Obtained using the YouTube API & Python

⇢ Program takes in keywords & outputs the total number of views for videos under that keyword
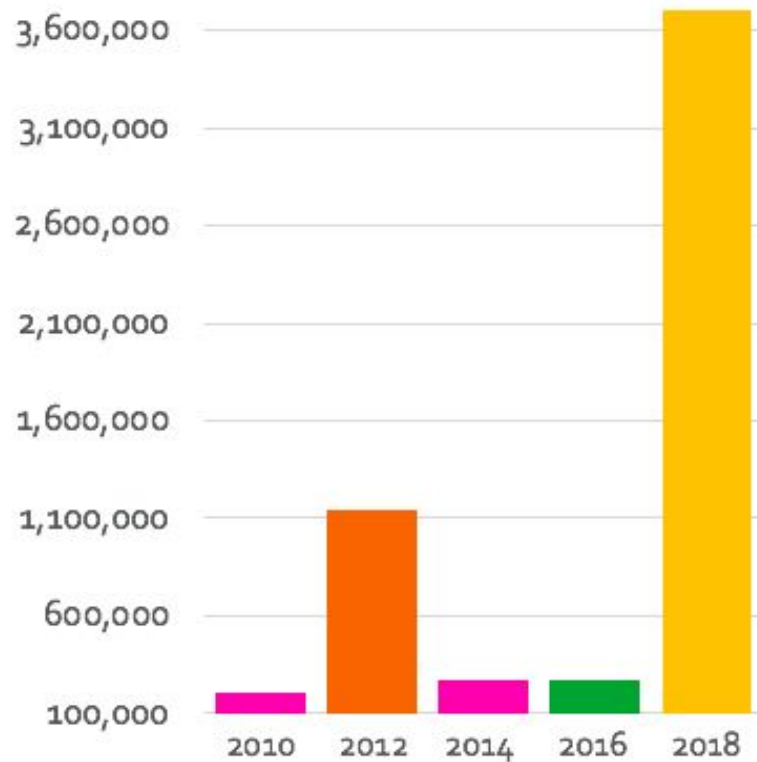
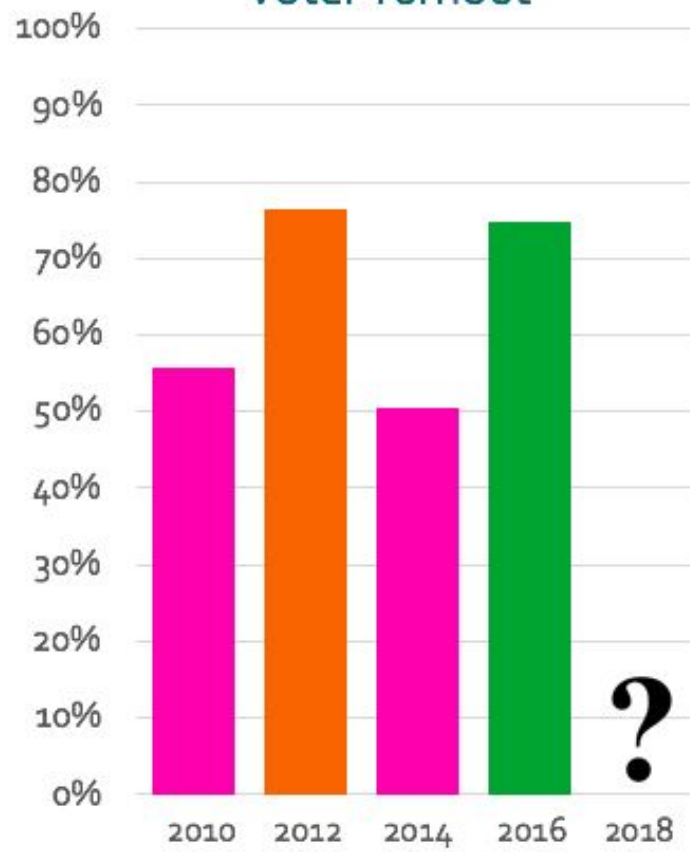# Correlation between YouTube Views & Voter Turnout

Keywords

- "Minnesota democrat"
- "Minnesota republican"
- "Minnesota election"
- "Minnesota general"
- "Minnesota primary"

- "Minnesota house of representatives"
- "Minnesota senate"
- "Minnesota vote"
- "Minnesota voter registration

Correlation between YouTube Views & Voter Turnout

Total Views

Voter Turnout

Correlation Coefficient

--→ 0.626

Correlation between YouTube Views & Voter Turnout

Most Viewed Keywords

→ 2010: "Minnesota governor"

→ 2012: "Michele Bachmann"

→ 2014: "Jack Shepard"

→ 2016: "Minnesota Clinton"

→ 2018: "Amy Klobuchar"

Correlation between YouTube Views & Voter Turnout

Explanations?

⤏ 2012: Michele Bachmann → former US House of Representative for MN; candidate for the Republican nominee for the 2012 presidential election

⤏ 2014: Jack Shepard → faced arson charges for burning down his own dentist office & was convicted of various crimes; ran for US Senate against Al Franken in 2014

⤏ 2018: Amy Klobuchar → MN senator; confronted during Brett Kavanaugh during his Senate confirmation hearing