

---

# *As Ohio goes, so goes the nation:*

## Predictors of U.S. presidential elections

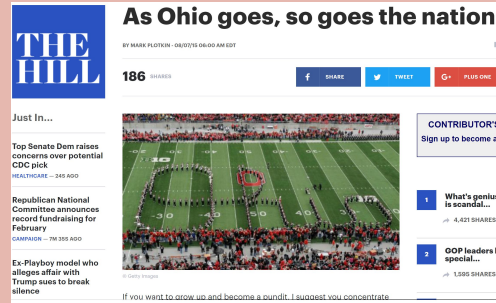
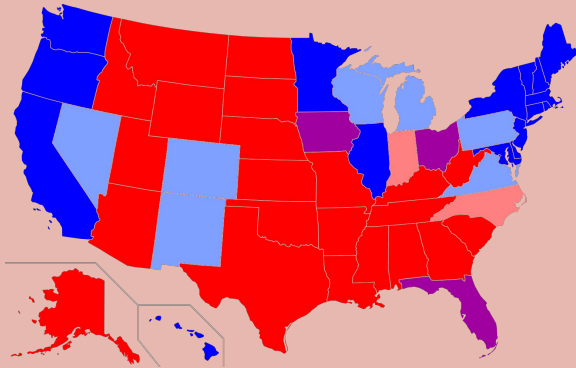
Juan M. Alzate Vanegas  
University of Central Florida  
3-20-2018





A common belief in U.S. politics is that there are **certain states** that presidential candidates must carry in order to win the general election.

These are thought to be the **swing states** or “**purple**” states, competitive states that favor neither major party by a large margin.



According to *FiveThirtyEight*:

#### Traditional swing states

- Colorado
- Florida
- Iowa
- Michigan
- Minnesota
- Nevada
- New Hampshire
- North Carolina
- Ohio
- Pennsylvania
- Virginia
- Wisconsin

#### Proposed swing states

- Arizona
- Georgia
- Maine
- Utah

"No Republican has ever been elected president without winning Ohio."

IVs	DVs
<i>Election year</i>	Popular vote
<i>State</i>	Result of each state and nationwide election
<i>Candidate party</i>	Whether state election winners go on to win the presidency
<i>Candidate</i>	

## Questions of interest

1. *How has **voter turnout** changed over time?*
2. *Are certain parties **more likely to win** certain elections than other parties?*
3. *Is there an **association** between the outcome of a state election and the outcome of the presidential election?*
4. *Do these effects **interact**?*

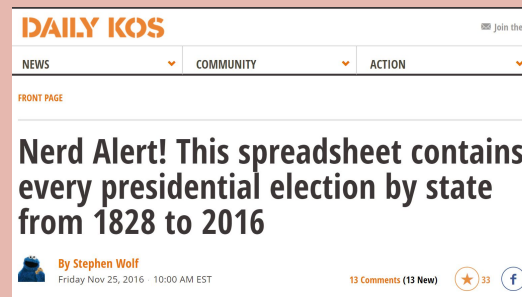
# Dataset: *RawElectionData.csv*

Year	Candidate	Nationwide	Alabama	Alaska	Arizona
2016	Total	137,098,601	2,123,372	318,608	2,604,657
2016	Hillary - Dem	65853625	729547	116454	1161167
2016	Donald - Rep	62985106	1318255	163387	1252401
2016	Clinton, Gary - Lib	4489233	44467	18725	106327
2012	Total	129235558	2074338	300495	2306559
2012	Barack - Dem	65918507	795696	122640	1025232

Includes...

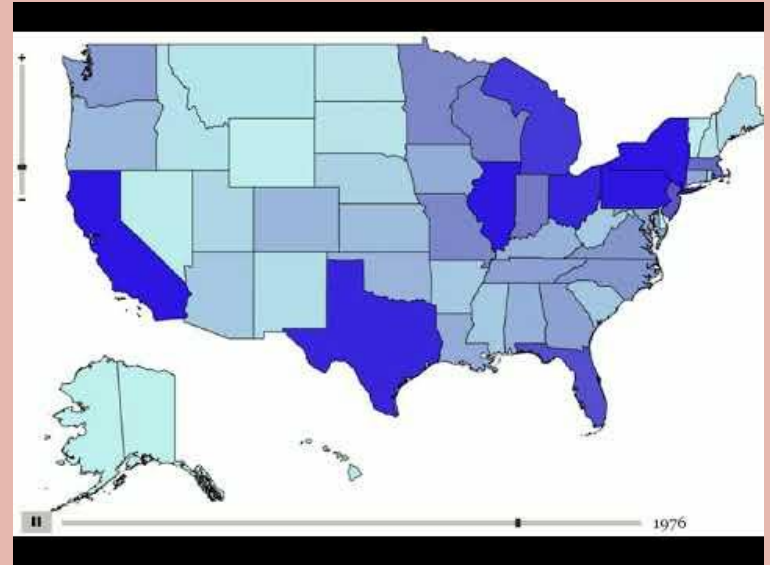
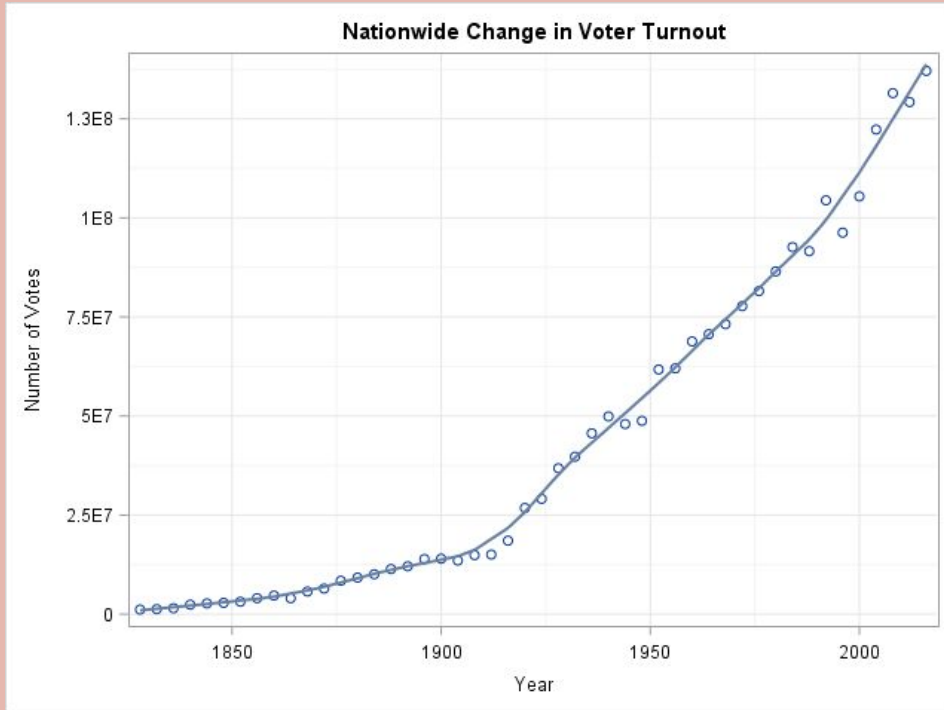
- Popular vote for each state and nationwide presidential election from 1828\*-2016
- Only the top 3 contenders or “other”

\*Certain current states (e.g., Alaska, Hawaii, etc.) did not hold elections during a portion of this time. Thus, the historical data are incomplete.



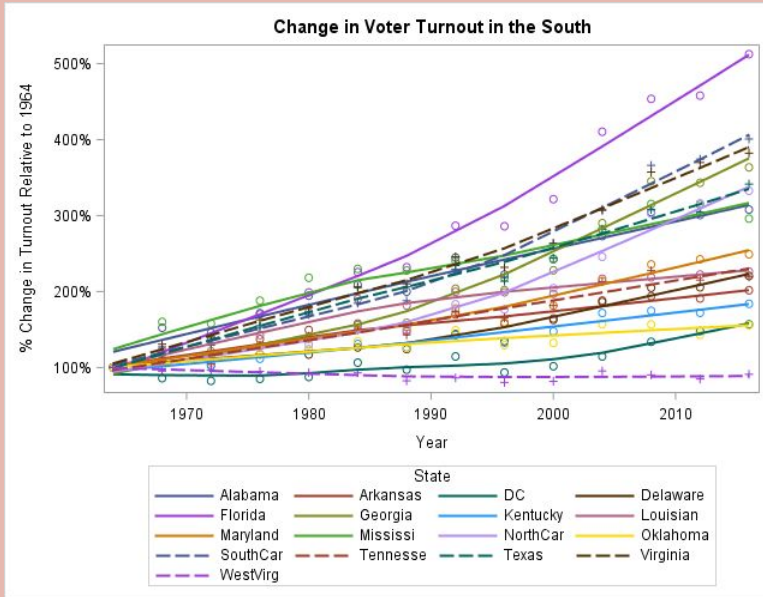
Election Results				
Year	Party	Candidate	State	Vote
2016	TOT	Total	Nationwi	137098601
2016	TOT	Total	Alabama	2123372
2016	TOT	Total	Alaska	318608
2016	TOT	Total	Arizona	2604657
2016	TOT	Total	Arkansas	1130635
2016	TOT	Total	Californ	14237884

# 1. How has **voter turnout** changed over time since the first election recorded in the dataset?

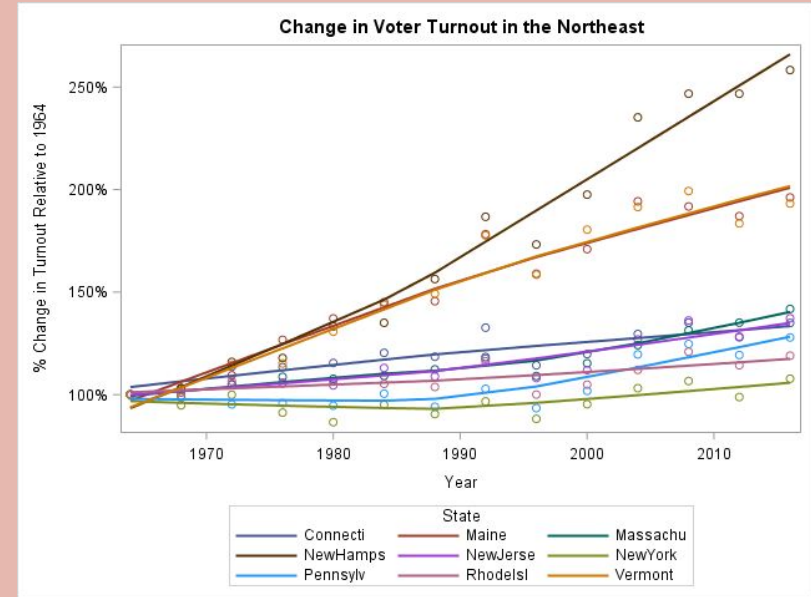


(Made with SAS and OpenHeatMap)

# By U.S. Census Region

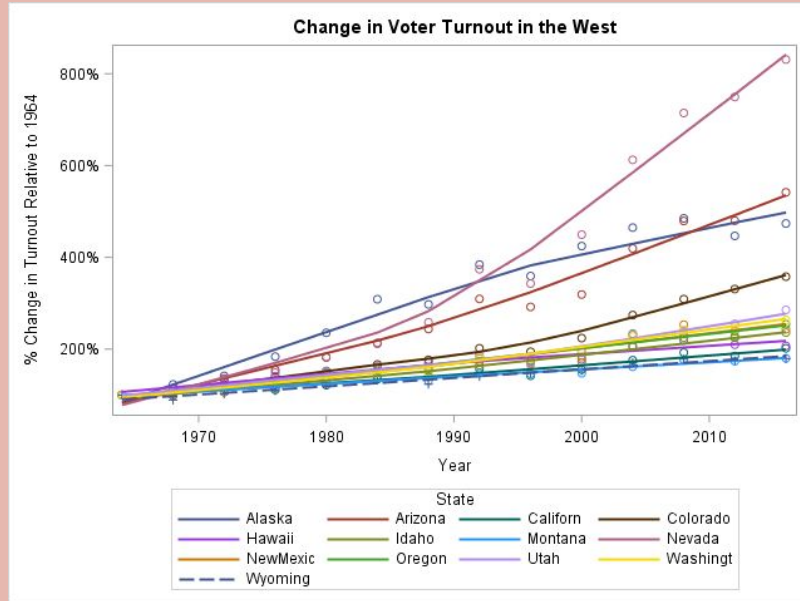


**South**

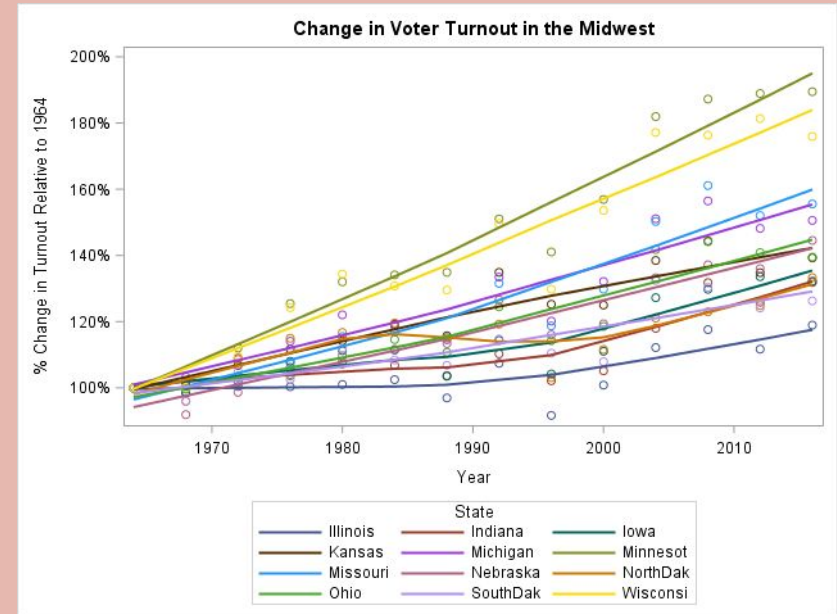


**Northeast**

# By U.S. Census Region



**West**



**Midwest**

## 2. Are certain parties *more likely to win* certain elections than other parties?

- 19 parties (plus Independents & Other) are represented in the dataset, but most are not frequently observed.
- Here, I only consider the subset of **Democrats and Republicans**.

### Conclusions:

- Democrats won **45.74%** (918/2,007) of state elections in which they competed.
- Republicans won **55.03%** (1,018/1,850) of state elections in which they competed.

**Fisher's exact test** suggests this difference is statistically significant ( $p < .001$ ).

Election Data

The FREQ Procedure

Frequency Percent Row Pct Col Pct	Table of Party by Winner			
	Party	Winner		Total
		0	1	
DEM		1089	918	2007
		28.23	23.80	52.04
		54.26	45.74	
		56.69	47.42	
REP		832	1018	1850
		21.57	26.39	47.96
		44.97	55.03	
		43.31	52.58	
Total		1921	1936	3857
		49.81	50.19	100.00

Statistics for Table of Party by Winner

Statistic	DF	Value	Prob
Chi-Square	1	33.2122	<.0001
Likelihood Ratio Chi-Square	1	33.2610	<.0001
Continuity Adj. Chi-Square	1	32.8418	<.0001
Mantel-Haenszel Chi-Square	1	33.2036	<.0001
Phi Coefficient		0.0928	
Contingency Coefficient		0.0924	
Cramer's V		0.0928	

Fisher's Exact Test

Cell (1,1) Frequency (F)	1089
Left-sided Pr <= F	1.0000
Right-sided Pr >= F	<.0001
Table Probability (P)	
Two-sided Pr <= P	<.0001



### 3. Is there an **association** between the outcome of a state election and the outcome of the presidential election?

In other words, can winning certain state elections predict the winner of the general election?

Statistics for Table of State by PredictiveOfWinner

Statistic	DF	Value	Prob
Chi-Square	50	67.0578	0.0539
Likelihood Ratio Chi-Square	50	67.8206	0.0474
Mantel-Haenszel Chi-Square	1	0.5039	0.4778
Phi Coefficient		0.1815	
Contingency Coefficient		0.1786	
Cramer's V		0.1815	

Sample Size = 2036

Conclusion: **Yes.**

The **Likelihood Ratio Test (LRT)** suggests this difference is statistically significant at the  $\alpha = 5\%$  level,  $G^2(50) = 67.82$ ,  $p = .0474$ .

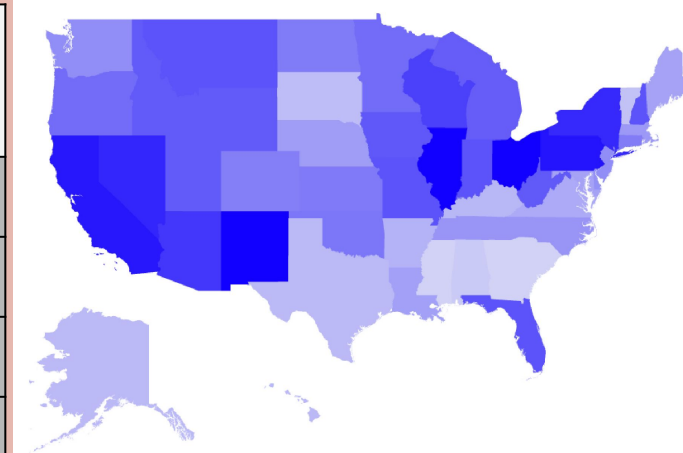
- The most predictive state?  
(place your bets!)

?

# Predictability by State



Most predictive elections	% of elections carried by president-elect	Least predictive elections	% of elections carried by president-elect
1. New Mexico	88.89% (24/27)	1. D.C.	42.86% (6/14)
2.5. Ohio	83.33% (40/48)	2. Mississippi	53.33% (24/45)
2.5. Illinois		3. Georgia	54.34% (25/46)
4. Pennsylvania	81.25% (39/48)	4. South Carolina	55.26% (39/48)
5. California	80.95% (34/42)	5. Alabama	56.82% (25/44)



(Made with SAS and OpenHeatMap)

#### 4. Are there any *interaction* effects between State and Party?

Using **Probit regression** to predict whether the winner of a state election wins the presidency, the following model was created to study the effects of State and Party on predictability within the Democrat/Republican subset:

$$p(\text{Win presidency} \mid \text{State, Party}) = \Phi(\alpha + \beta(\text{State}) + \beta(\text{Party}) + \beta(\text{State} \times \text{Party}))$$

- Model: the model was an adequate fit for the data.
- Interaction: several interaction effects were statistically significant.
- Marginal effects
  - State: marginally significant,  $\chi^2(50) = 65.94$ ,  $p = 0.0648$ .
    - Removed from model.
  - Party (Dem vs. Rep, Rep fixed):  $\beta = -1.56$ ,  $\chi^2(1) = 7.02$ ,  $p < .01$ .

Criteria For Assessing Goodness Of Fit			
Criterion	DF	Value	Value/DF
Log Likelihood		-1024.5367	
Full Log Likelihood		-1024.5367	
AIC (smaller is better)		2253.0734	
AICC (smaller is better)		2264.2560	
BIC (smaller is better)		2823.4433	

Algorithm converged.

**Democrats were significantly more likely\* to win the presidency when they carried...**

Arkansas ( $\beta = 1.28$ )	New York ( $\beta = 1.10$ )
Florida ( $\beta = 1.22$ )	North Carolina ( $\beta = 0.91$ )
Illinois ( $\beta = 1.10$ )	Virginia ( $\beta = 0.91$ )
Louisiana ( $\beta = 0.91$ )	
Missouri ( $\beta = 1.34$ )	
New Mexico ( $\beta = 1.47$ )	

**Republicans were significantly more likely\* to win the presidency when they carried...**

Alabama ( $\beta = -1.79$ )	Kentucky ( $\beta = -1.56$ )	North Carolina ( $\beta = -1.56$ )
Arkansas ( $\beta = -1.79$ )	Louisiana ( $\beta = -1.68$ )	South Carolina ( $\beta = -1.56$ )
Delaware ( $\beta = -1.25$ )	Maryland ( $\beta = -1.46$ )	Tennessee ( $\beta = -1.33$ )
Florida ( $\beta = -1.26$ )	Mississippi ( $\beta = -1.87$ )	Texas ( $\beta = -1.62$ )
Georgia ( $\beta = -2.17$ )	Missouri ( $\beta = -1.13$ )	Virginia ( $\beta = -1.56$ )
Hawaii ( $\beta = -2.24$ )	New Jersey ( $\beta = -1.25$ )	Washington ( $\beta = -1.19$ )

\* $p < .05$

# Discussion

- Importance and usefulness of this analysis
  - Reasonably comprehensive analysis
    - $N = 8,996$  observations
    - 1828-2016
    - 50 States + D.C.
    - 19 parties
  - Electoral campaign strategy: plan to visit (and win) predictive states.
  - Predicting or estimate campaign outcomes before elections.
  - Related variables can be added to the model to assess other predictors of election outcomes.
- Ideas for related projects
  - Changes in voter turnout over time relative to changes in population.
  - Choose specific time periods/regions of focus as subsets.
    - e.g., predictability by region, predictability in 1828 vs. now, etc.
  - Compare swing states vs. non-swing states different definitions of swing states, etc.
  - Factors related to campaign finance, advertising, electoral college votes, etc.



# Works Cited

- Dataset
  - <https://dailykos.com/stories/2016/11/25/1601042/-Nerd-Alert-This-spreadsheet-contains-every-presidential-election-by-state-from-1828-to-2016>
- Works consulted
  - <https://fivethirtyeight.com/features/the-odds-of-an-electoral-college-popular-vote-split-are-increasing/>
  - [https://web.archive.org/web/20130921053705/http://www.census.gov/geo/maps-data/maps/pdfs/reference/us\\_regdiv.pdf](https://web.archive.org/web/20130921053705/http://www.census.gov/geo/maps-data/maps/pdfs/reference/us_regdiv.pdf)
  - <http://www.openheatmap.com>