Presidential Campaigns

STEPHEN ALLEN, JACE BARTON, LEON CHEN, ANDREW REUBEN,
SAMRUDDHI SOMANI

Agenda

Question

Context

Data Preprocessing

Modeling

Takeaways

Further Research

Appendix

Question

In which states should ad money be spent?

Are these the states that ad money is spent in?

Context

Candidate has to win 270 of 538 electoral votes

- Two to each state
- Three to DC
- Balance allocated among states by population

Every state except Nebraska and Maine is “winner-take-all”

Data Sources

Bureau of Economic Analysis

United States Census Bureau

American National Elections Studies

The American Presidency Project

Fair Vote

Elect Project

Data Preprocessing

ISSUE	*SOLUTION*
Limited data	Limit analysis to 2008 and 2012
P>>N	Remove duplicates, PLSR
Fractional Shares	Logit Transform
Interested in ‘swinginess’	Target=$a-a^2$

Modeling Considerations

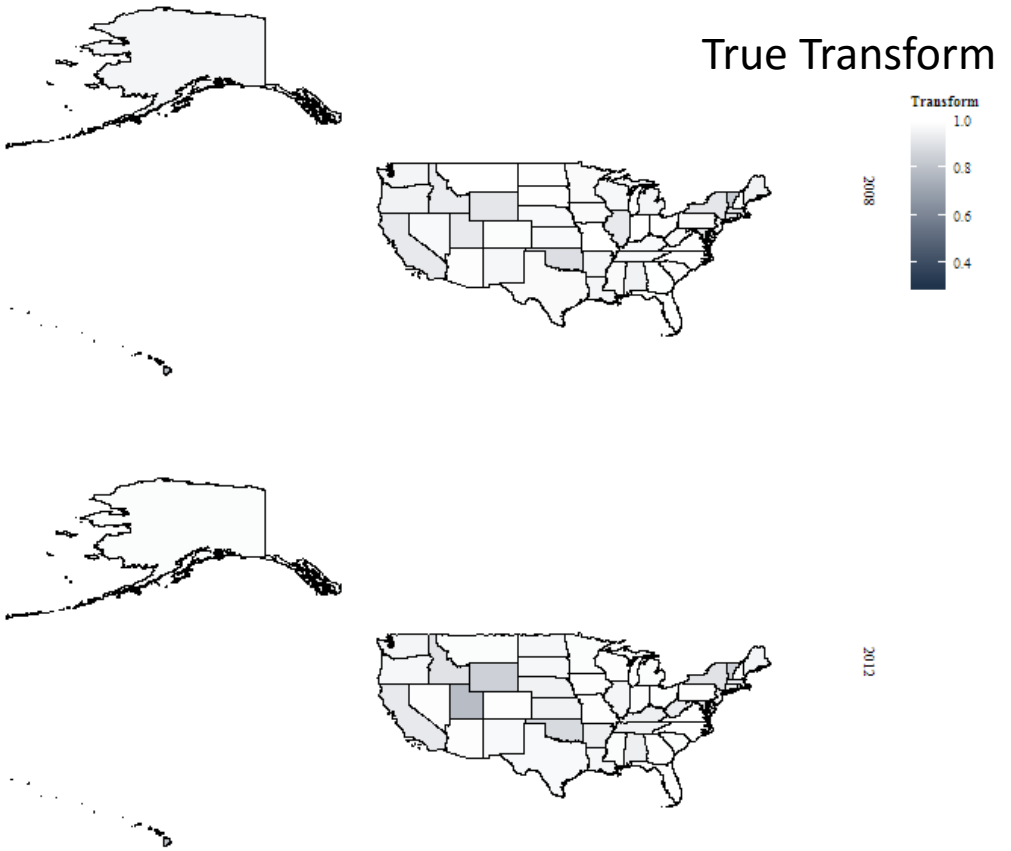
Electoral votes don’t change; swinginess does

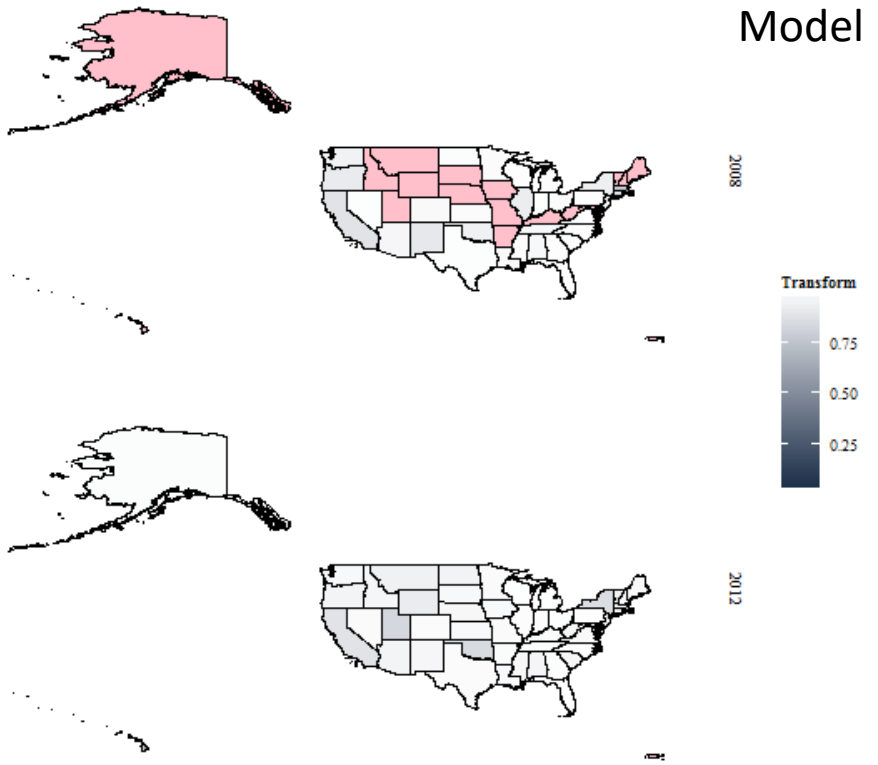
Model swinginess first

Then model effects of swinginess and electoral votes on ad spend.

Use smoothing factors to handle logarithms.

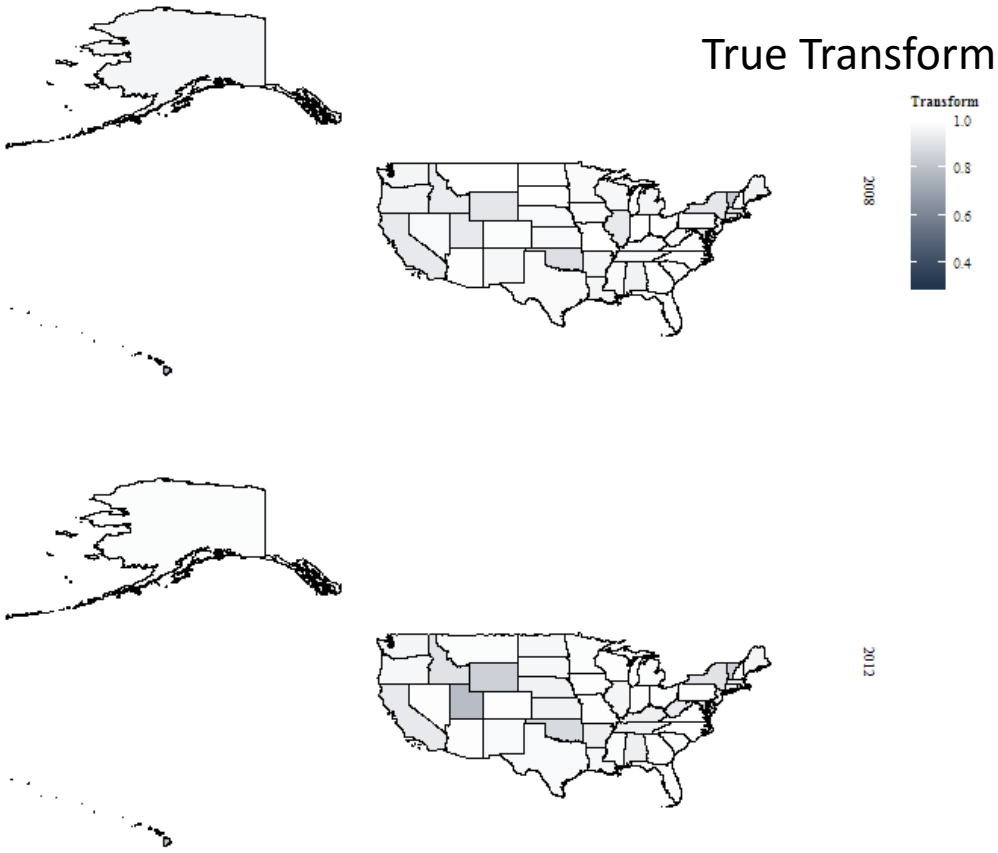
Demographics (RMSE=0.0875)



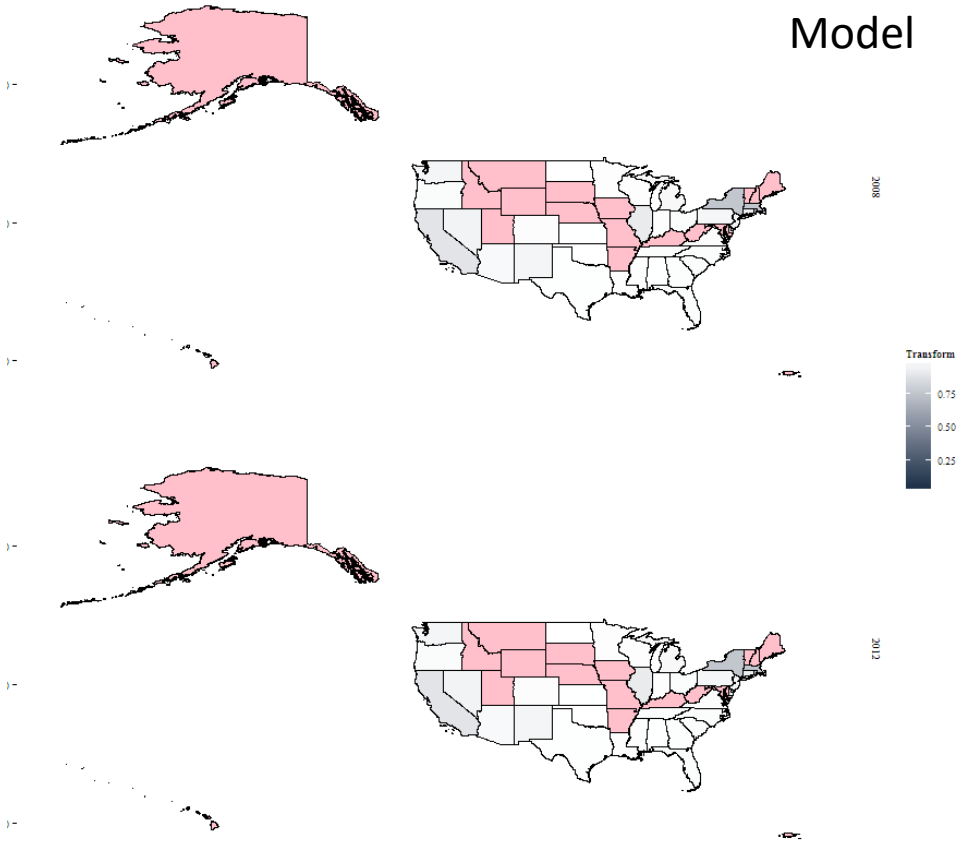


Past Behavior (RMSE=0.0731)

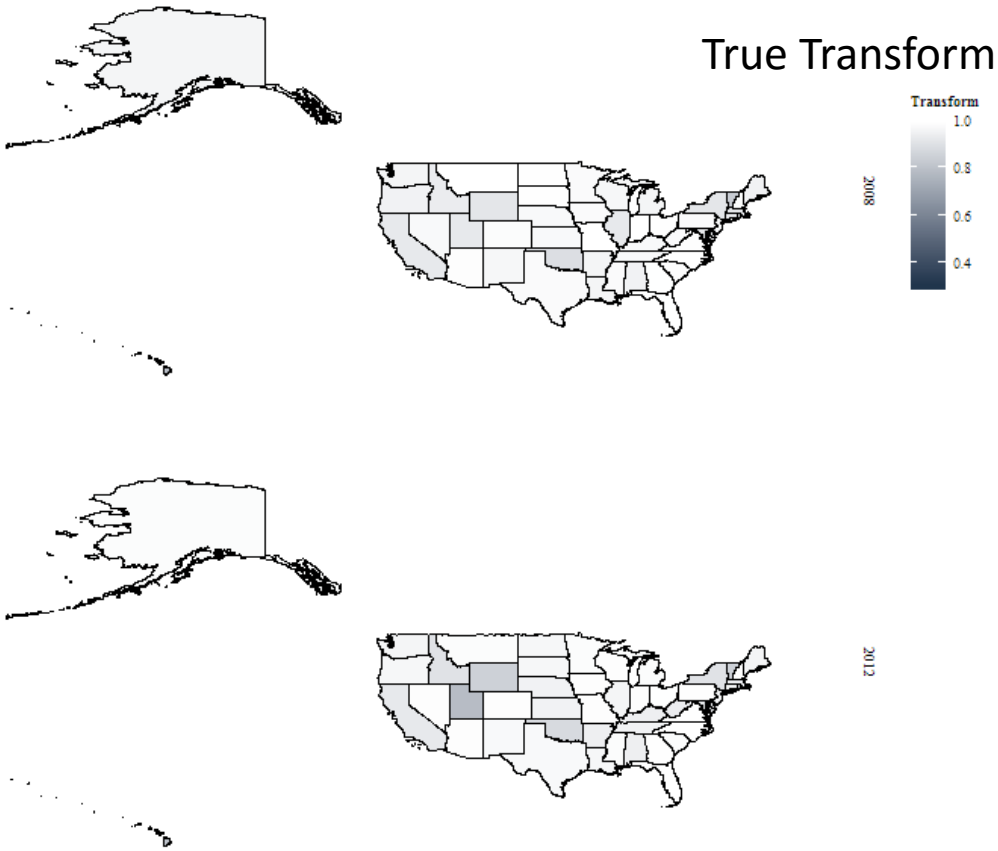
True Transform

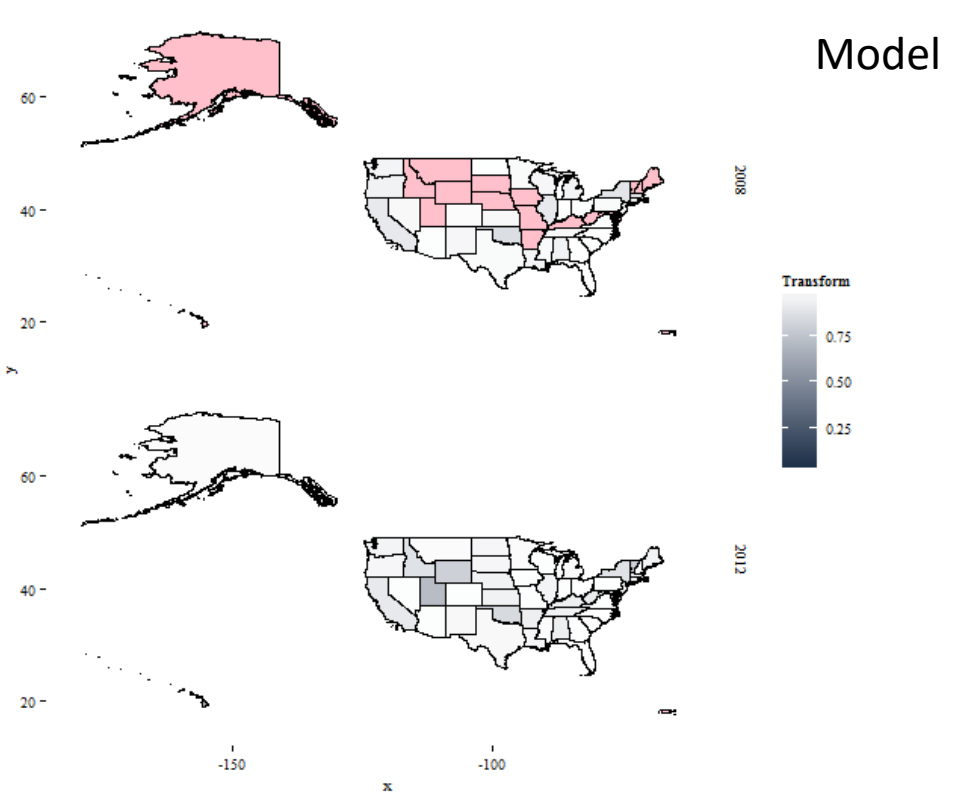


Model



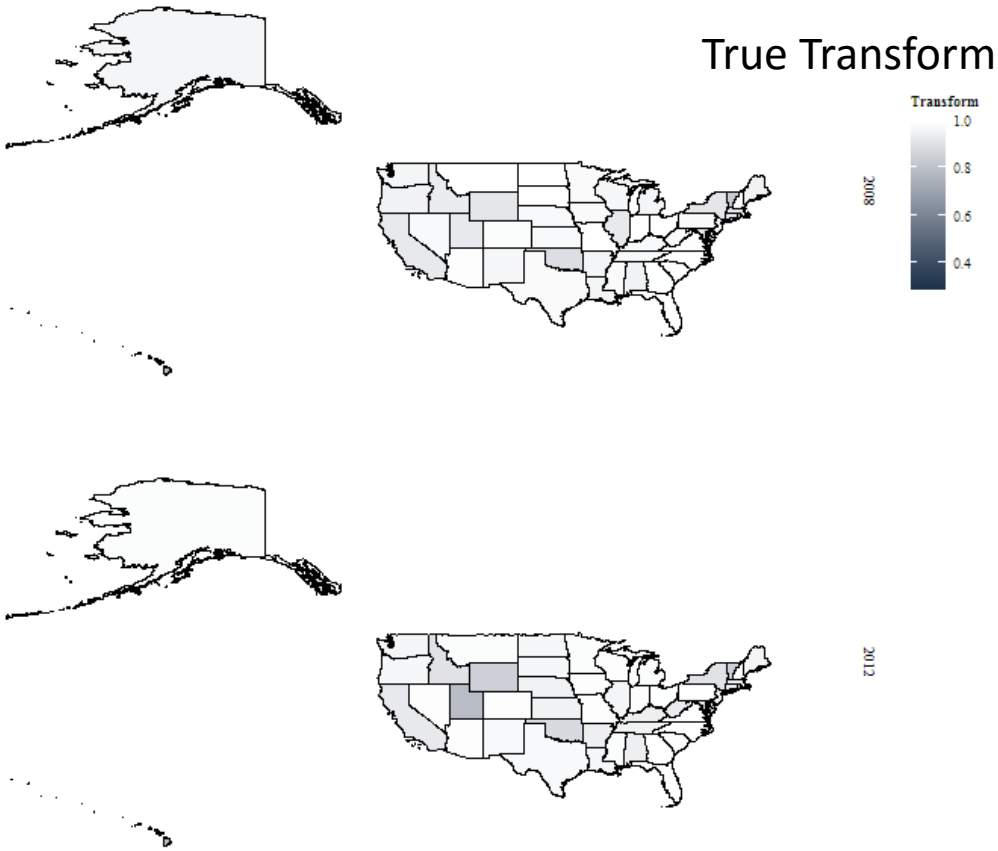
Attitudes (RMSE=.0445)



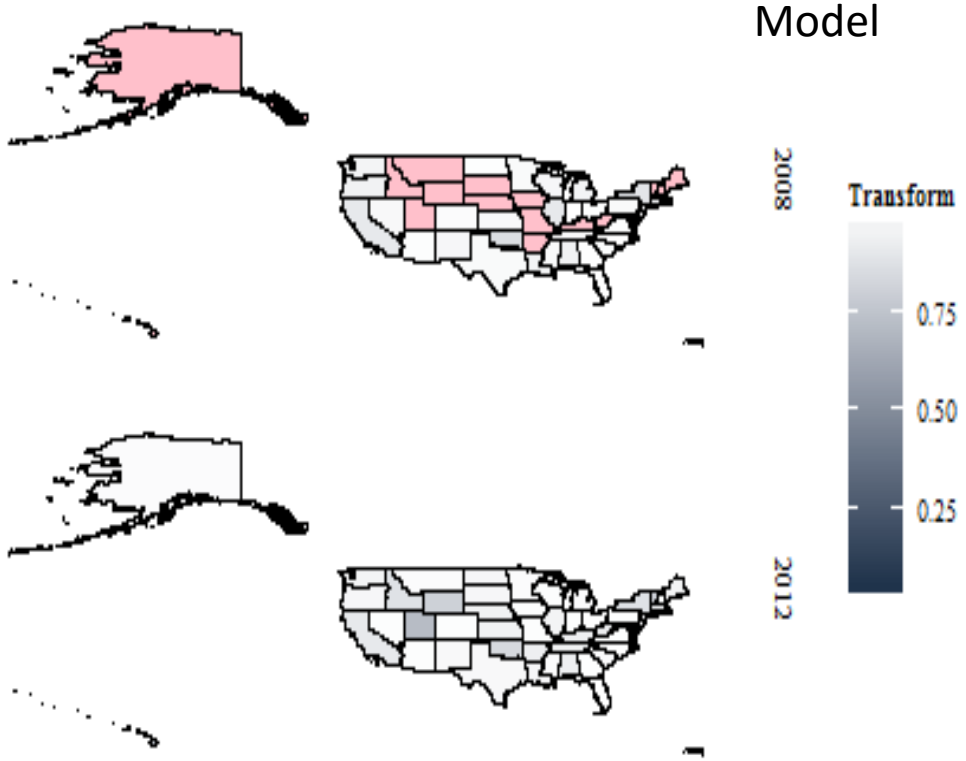


Everything (RMSE=.0444)

True Transform



Model



Ad Spending?

Suitable model for swinginess (but overestimates swinginess)

Spearman correlation against swinginess: 62%

Spearman correlation against electoral votes: 31%

Spearman correlation against interaction: 99%

- Clearly, both matter

What should you include in your ads?

Most and Least Controversial Groups

LEAST CONTROVERSIAL	*MOST CONTROVERSIAL*
Middle Class	Illegal Immigrants
Young People	LGBT
White People	Christian Fundamentalists
Military	Republican Party
Southerners	Unions
Poor	Democratic Party

Most and Least Popular Groups

MOST POPULAR

Military

Middle Class

Young

White

Southern

Poor

LEAST POPULAR

Republican Party

Illegal Aliens

Congress

Federal Government

Takeaways

Spend in big swingy states

Past behavior is predictive and easily obtained

Attitude is more expensive but more predictive

- Limited collection might be an appropriate solution.

Phrasing and issues matter

Further Research

Examine at county level

Examine for midterm elections

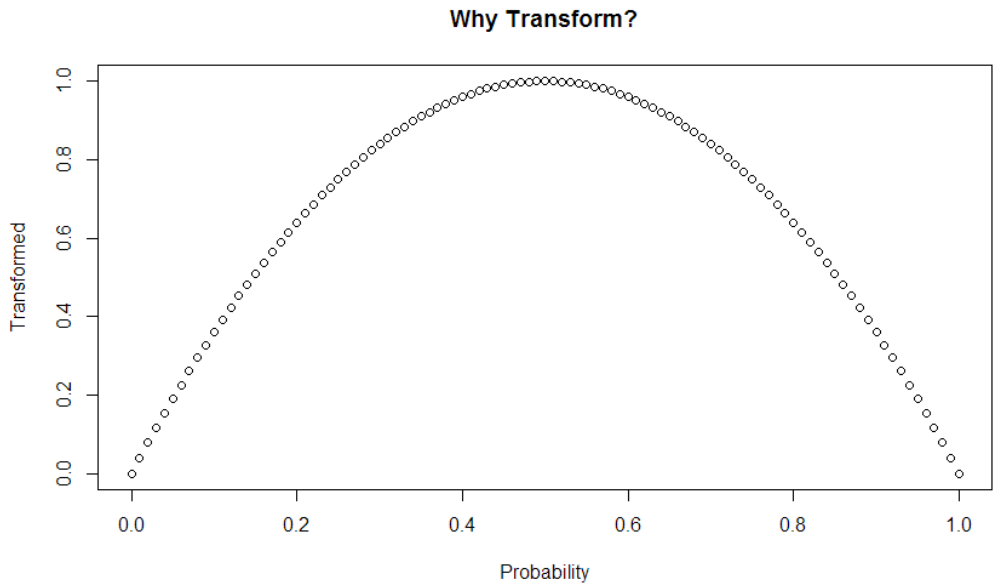
Examine social media for attitude data

Examine model predictions against 2016 results

Appendix

Data Processing Steps

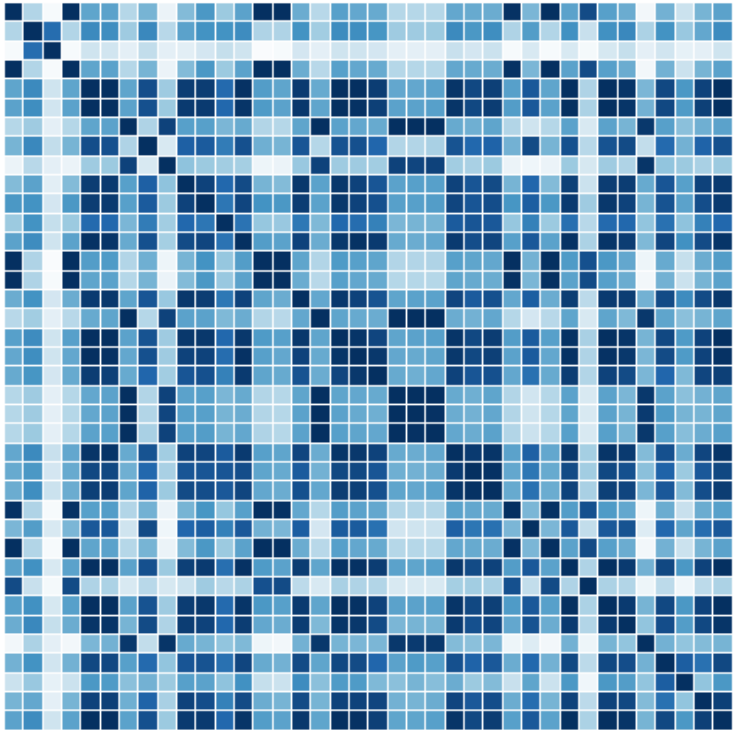
Why the quadratic transform?



Which side of 50% is irrelevant for swinginess.

We care how far away for 50%

Why Partial Least Squares?



The darker the square, the more correlated.

Rampant multicollinearity