

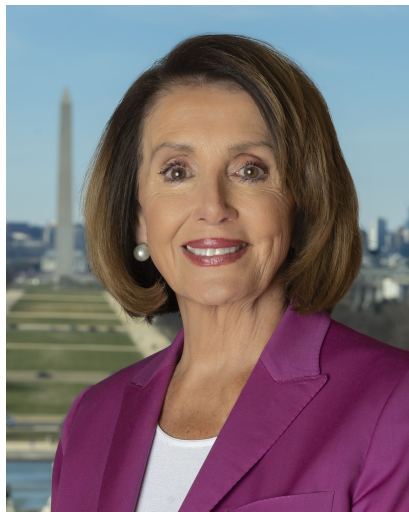
Political Subgroup Visualization

By: Nathan Lee



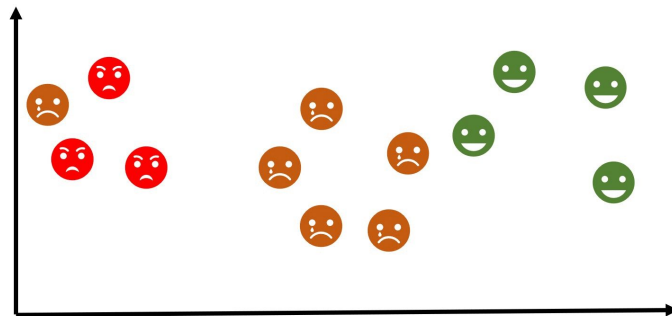
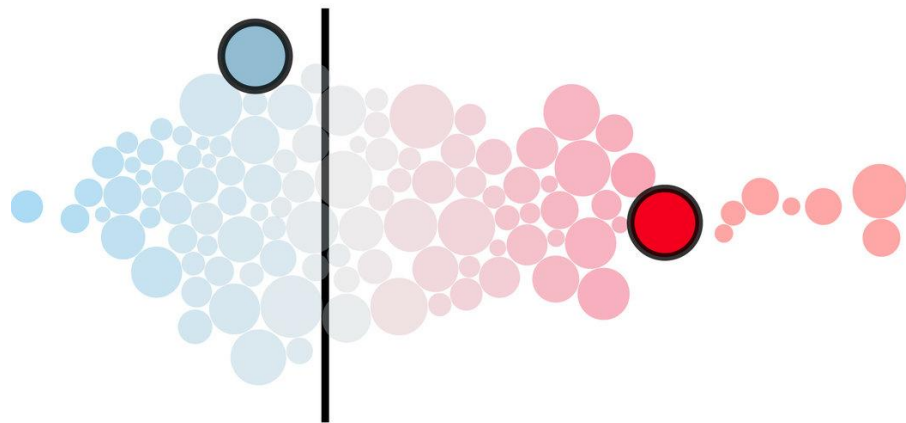


Problem





Goal





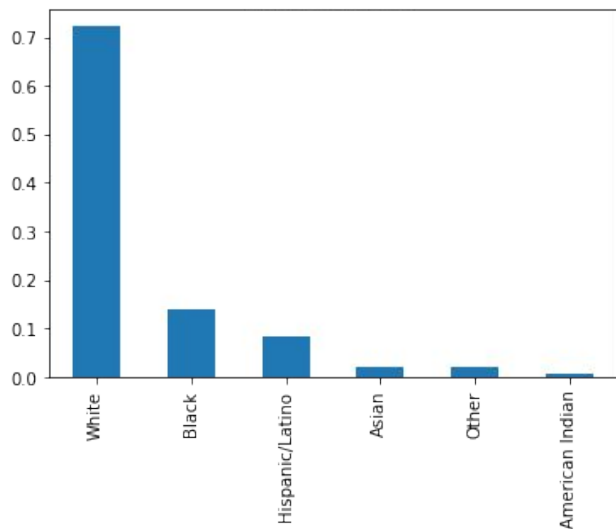
Data

- National Election Exit Poll Data: 2018 Midterms
- Edison Research
- November 6, 2018
- Sample: Existing voters, new voters, and early/absentee voters
- Size: 19221--> 14305
- Total Number of Uniform Features: 20
- Total Number of Unique Features: 48

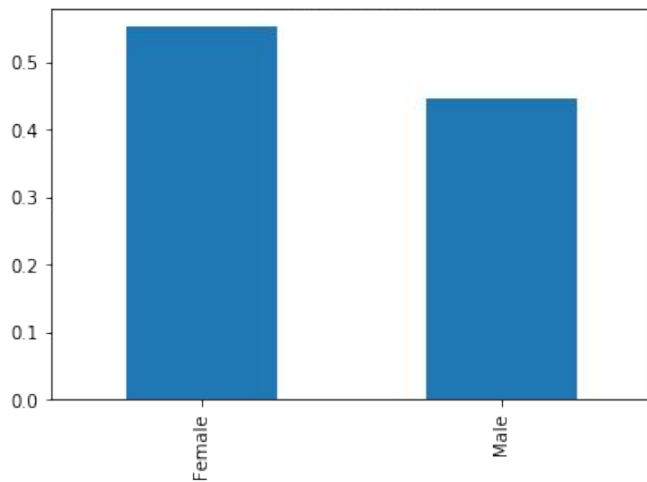


Demographic Breakdown

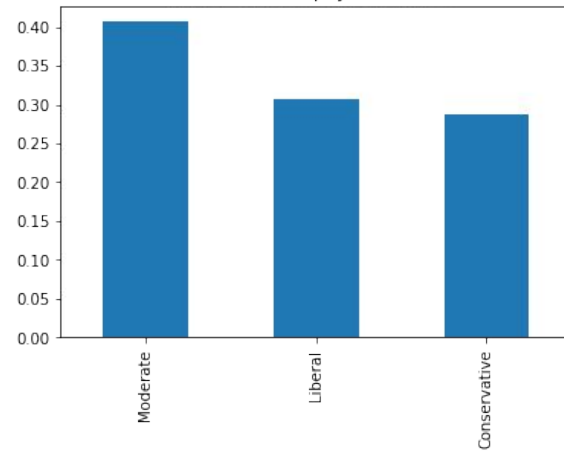
Race Breakdown



Sex Breakdown



Political Philosophy Breakdown



Issues and Methodology

[A] Are you:

- ☐ Male **SEX**
☐ Female

[B] Are you:

- ☐ White **QRACEAI**
☐ Black
☐ Hispanic/Latino
☐ Asian
☐ American Indian
☐ Other

[C] In today's election for U.S. Senate, did you just vote for:

- ☐ The Democratic candidate
☐ The Republican candidate
☐ Other: Who? _____
☐ Did not vote **SENU18**

[D] In today's election for governor, did you just vote for:

- ☐ The Democratic candidate
☐ The Republican candidate
☐ Other: Who? _____
☐ Did not vote **GOVUS18**

[E] Are you of Hispanic or Latino descent?

- ☐ Yes
☐ No **LATINO**

[F] In which age group are you?

- ☐ 18-24 ☐ 45-49
☐ 25-29 ☐ 50-59
☐ 30-39 ☐ 60-64 **AGE8**
☐ 40-44 ☐ 65 or over

[G] In today's election for U.S. House of Representatives, did you just vote for:

- ☐ The Democratic candidate
☐ The Republican candidate
☐ Other: Who? _____
☐ Did not vote **HOUUS18**

[H] Which best describes your education? You have:

- ☐ Never attended college
☐ Attended college but received no degree **EDUC18**
☐ Associate's degree (AA or AS)
☐ Bachelor's degree (BA or BS)
☐ An advanced degree after a bachelor's degree (such as JD, MA, MBA, MD, PhD)

[I] In your vote for U.S. House, how would you rate the importance of _____?

- ☐ The most important factor
☐ An important factor
☐ A minor factor **BREAK18**
☐ Not a factor at all

[J] How do you feel about the way Donald Trump is handling his job as president?

- ☐ Strongly approve **TRUMP4**
☐ Somewhat approve
☐ Somewhat disapprove
☐ Strongly disapprove

[K] Have you ever served in the U.S. military?

- ☐ Yes **VETVOTER**
☐ No

[L] Is this the first time you have ever voted in a midterm election?

- ☐ Yes **FTVOTER18**
☐ No

[M] Is Donald Trump's approach to foreign policy making:

- ☐ The U.S. more safe
☐ The U.S. less safe
☐ No difference **TRUMPFP18**

[N] What should the Supreme Court do about Roe vs. Wade, the 1973 decision that made abortion legal?

- ☐ Keep it as is
☐ Overturn it **ABORTION18**

[O] In the 2016 election for president, did you vote for:

- ☐ Hillary Clinton (Dem)
☐ Donald Trump (Rep)
☐ Other **VOTE2016**
☐ Didn't vote

[P] Which concerns you more about today's election? That people who: **VOTEISS18**

- ☐ Should not be able to vote will cast ballots
☐ Should be able to vote will be prevented from voting

[Q] Would you rather see the U.S. House of Representatives controlled by: **CONTROLHOUSE**

- ☐ The Democratic Party
☐ The Republican Party

[R] How do you feel about Brett Kavanaugh's confirmation to the Supreme Court?

- ☐ Support **KAVANAUGH18**
☐ Oppose

[S] Do you think the government has done enough to protect this election from foreign interference? **ELECTSEC FOREIGN18**

- ☐ Yes ☐ No

[T] No matter how you voted today, do you usually think of yourself as a:

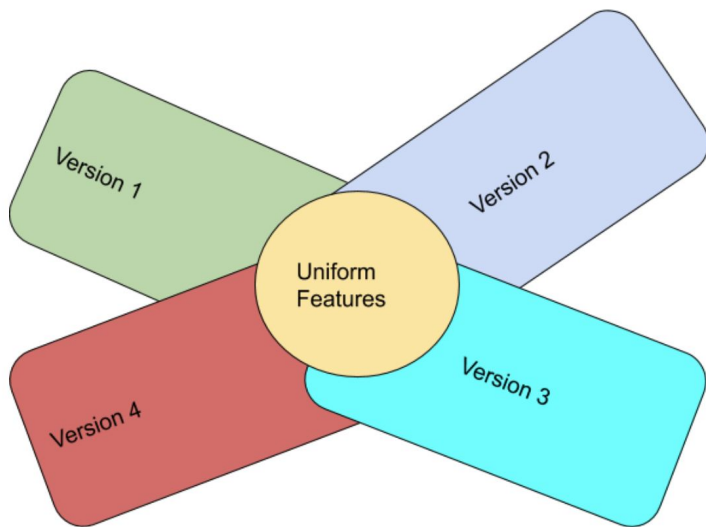
- ☐ Democrat
☐ Republican **PARTYID**
☐ Independent
☐ Something else

[U] On most political matters, do you consider yourself:

- ☐ Liberal **PHIL3**
☐ Moderate
☐ Conservative

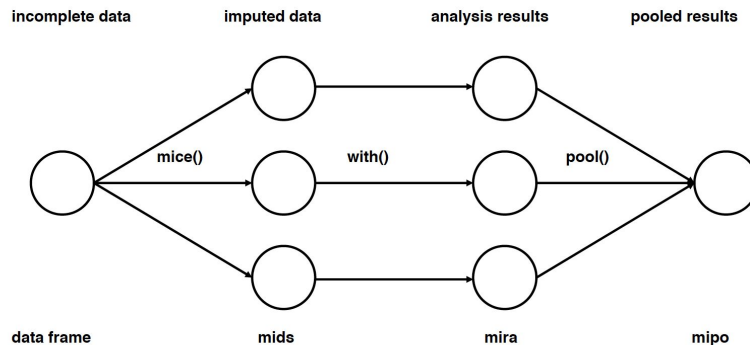


Data Imputation



Datawig

	col1	col2	col3	col4	col5			col1	col2	col3	col4	col5	
0	2	5.0	3.0	6	NaN	mean()		0	2.0	5.0	3.0	6.0	7.0
1	9	NaN	9.0	0	7.0			1	9.0	11.0	9.0	0.0	7.0
2	19	17.0	NaN	9	NaN			2	19.0	17.0	6.0	9.0	7.0





Datawig

#Initialize a SimpleImputer model

```
imputer = SimpleImputer(  
    input_columns=['title', 'text'],  
    output_column='finish',  
    output_path = 'imputer_model'  
)
```

#Fit an imputer model on the train data

```
imputer.fit(train_df=df_train)
```

```
predictions = imputer.predict(df_test)
```

#Calculate f1 score

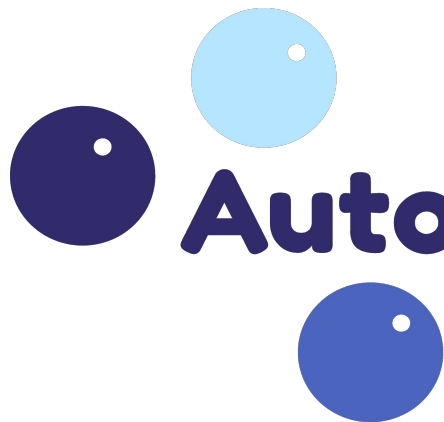
```
f1 = f1_score(predictions['finish'], predictions['finish_imputed'])
```

#Print overall classification report

```
print(classification_report(predictions['finish'], predictions['finish_imputed']))
```




Autoimputer



Autoimpute

Examine Patterns of Missing Data

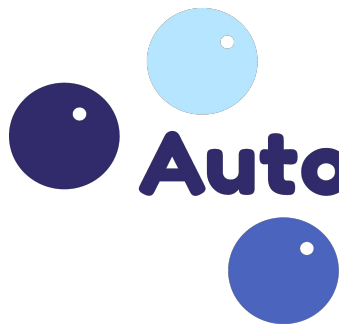
Innate handling of categorical variables as predictors and targets

Customizable data visualization tools

Numerous imputation methods



Autoimputer

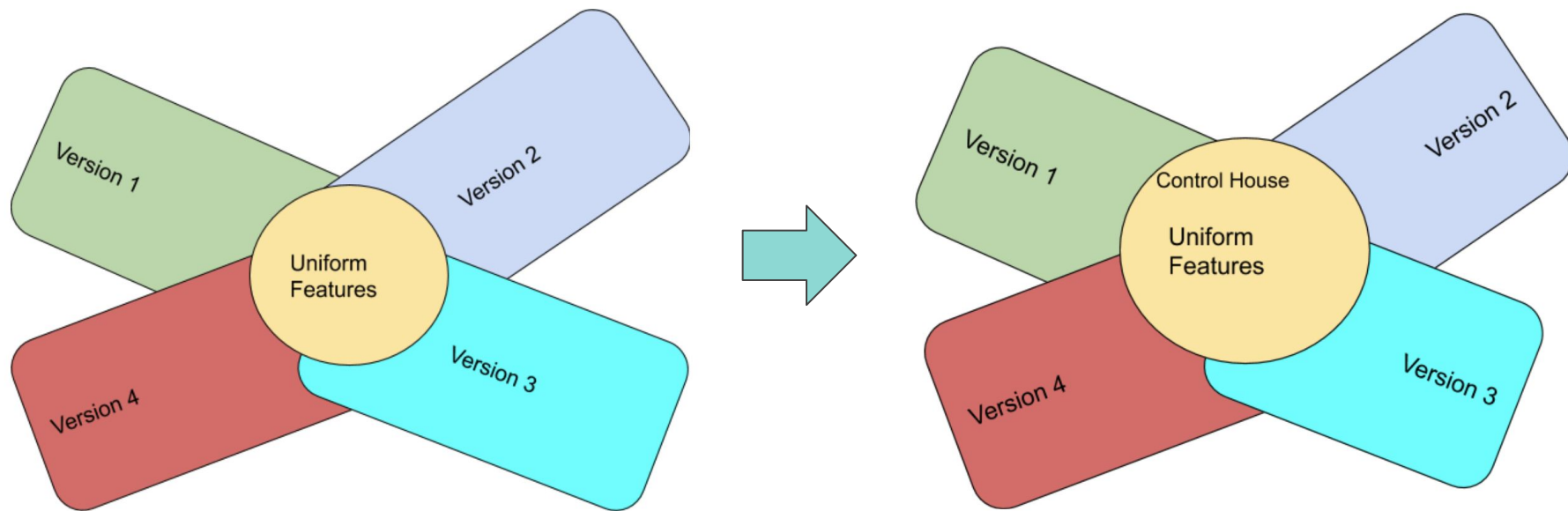


Numerous imputation
methods

Univariate	Multivariate	Time Series / Interpolation
Mean	Linear Regression	Linear
Median	Binomial Logistic Regression	Quadratic
Mode	Multinomial Logistic Regression	Cubic
Random	Stochastic Regression	Polynomial
Norm	Bayesian Linear Regression	Spline
Categorical	Bayesian Binary Logistic Regression	Time-weighted
	Predictive Mean Matching	Next Obs Carried Backward
	Local Residual Draws	Last Obs Carried Forward



Data Imputation Issues

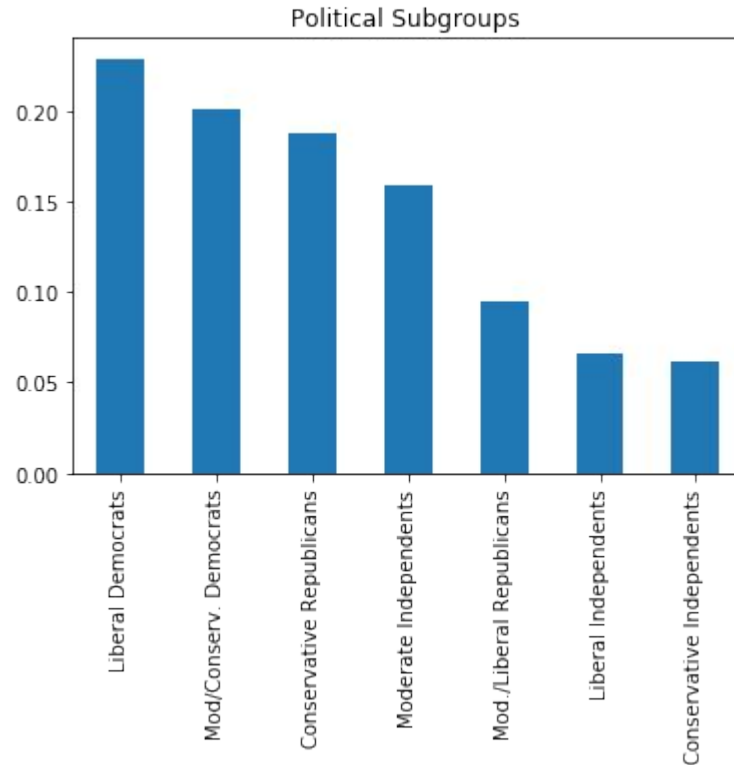
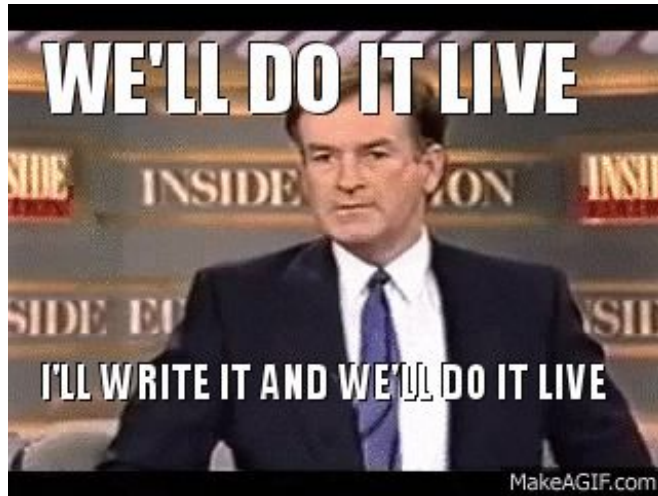


Unsupervised Algorithms

Algorithm	Silhouette Score
KMeans	$k = 2, \text{sil} = 0.1566$
DBSCAN	$\text{clusters} = 6, \text{sil} = 0.2422$
KModes	$\text{sil} = -0.2093$

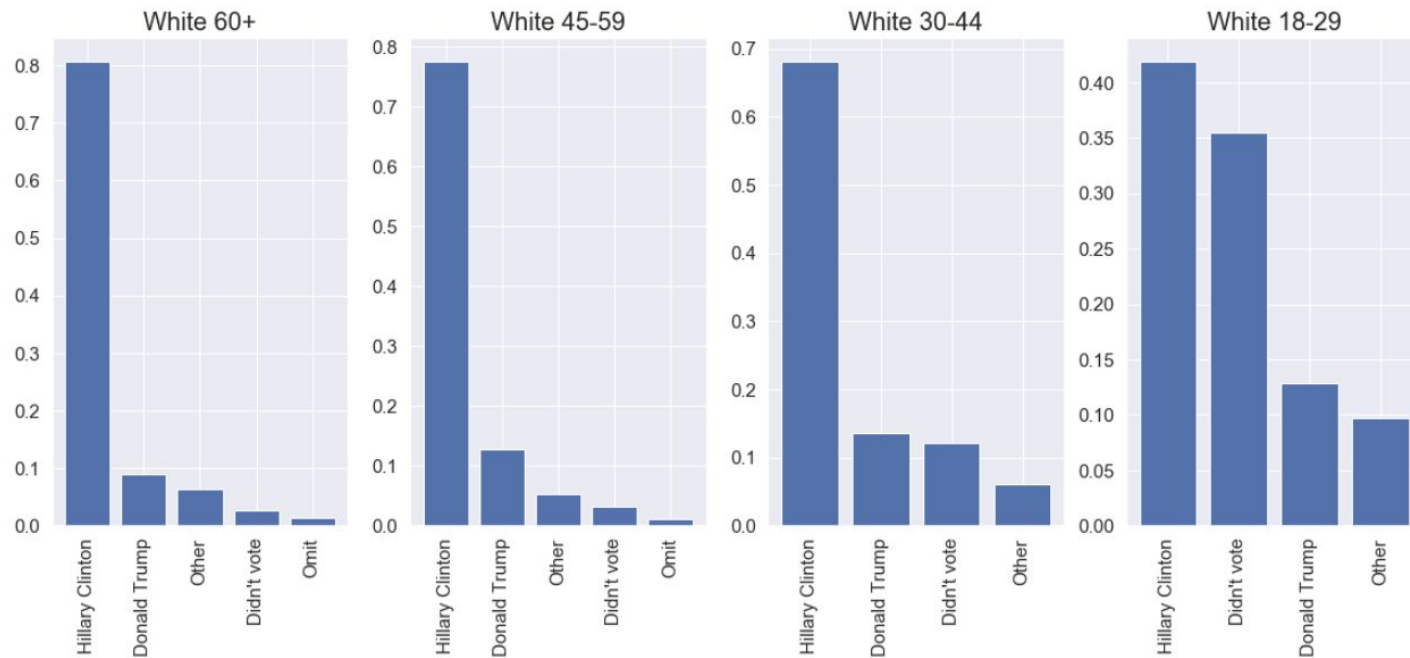


I guess this is what we're doing



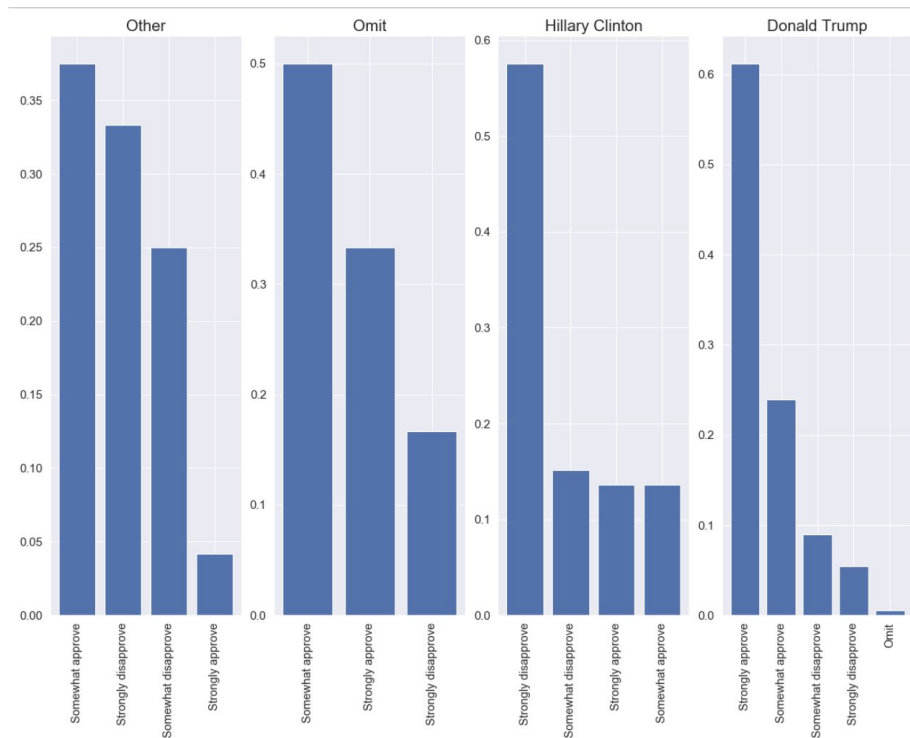
Version 1: Trump Approval, First Time Voters, and 2016 Election

Who the participant voted for in 2016 divided by age and race (mod/con Dems)

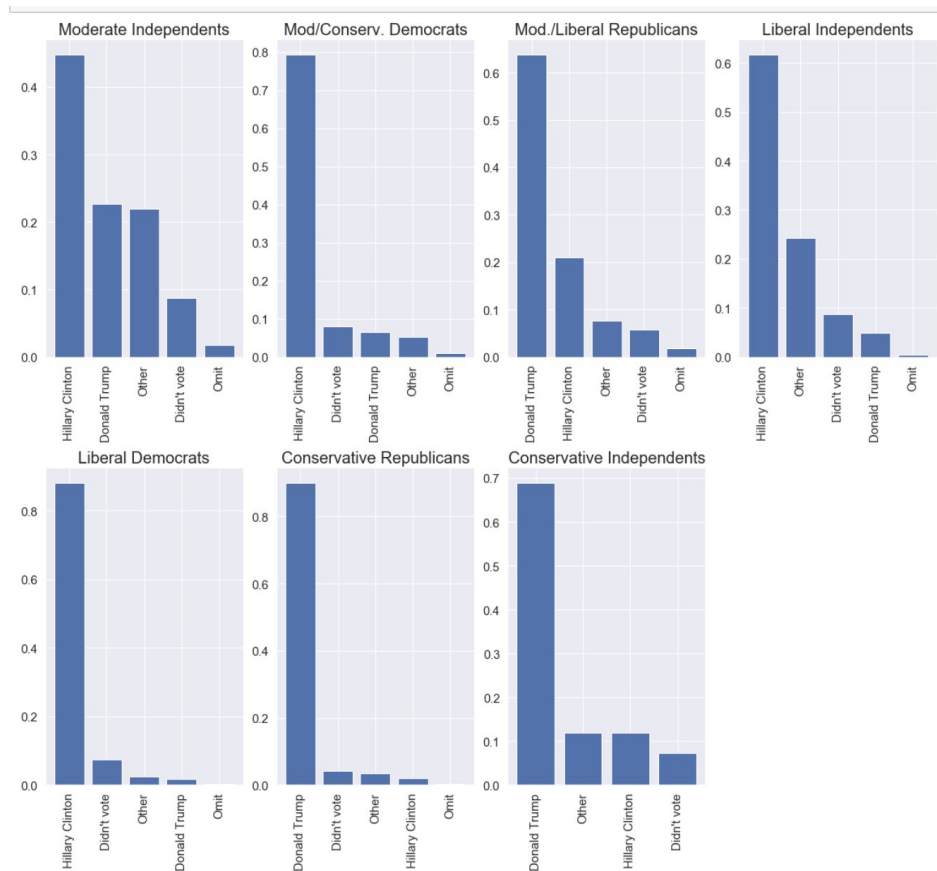


Version: 1 Continued

Mod/Liberal Republicans: Approval rating of Trump based on candidate participant voted for in 2016



Version 1: Continued



Version 2: Impeach, Russia Investigation, Trump Immigration, Race Factor, Issue Importance



[J] Which ONE of these four issues is the most important facing the country?
(CHECK ONLY ONE)

- 1 ☐ Health care
- 2 ☐ Immigration
- 3 ☐ The economy
- 4 ☐ Gun policy

[K] Is your opinion of Nancy Pelosi:

- 1 ☐ Favorable
- 2 ☐ Unfavorable

[L] How do you feel about Robert Mueller's handling of the investigation into Russian interference in the 2016 election?

- 1 ☐ Approve
- 2 ☐ Disapprove

[M] Are Donald Trump's immigration policies:

- 1 ☐ Too tough
- 2 ☐ Not tough enough
- 3 ☐ About right

[N] Should Congress impeach Donald Trump and remove him from office?

- 1 ☐ Yes
- 2 ☐ No

[O] How important is it that more racial and ethnic minorities be elected to public office?

- 1 ☐ Very important
- 2 ☐ Somewhat important
- 3 ☐ Not too important
- 4 ☐ Not at all important

[P] Do you think the investigation into Russian interference in the 2016 election is mostly:

- 1 ☐ Justified
- 2 ☐ Politically motivated

[Q] In general, how do you think things work in the U.S. today?

- 1 ☐ Whites are favored over minorities
- 2 ☐ Minorities are favored over whites
- 3 ☐ No group is favored

[R] Does the health care system in the United States need:

- 1 ☐ Major changes
- 2 ☐ Minor changes
- 3 ☐ No changes at all

[S] Are you:

- 1 ☐ Protestant
- 2 ☐ Catholic
- 3 ☐ Other Christian
- 4 ☐ Jewish
- 5 ☐ Muslim
- 6 ☐ Other
- 7 ☐ None

[T] No matter how you voted today, do you usually think of yourself as a:

- 1 ☐ Democrat
- 2 ☐ Republican
- 3 ☐ Independent
- 4 ☐ Something else

[U] On most political matters, do you consider yourself:

- 1 ☐ Liberal
- 2 ☐ Moderate
- 3 ☐ Conservative





Version 2



Impeachment
split on party
lines



Healthcare: Liberals
Economy: Conservatives



Russia
Investigation



Version 3: Women in Politics, Sexual Harassment, Parents Approval, Gun Control

Most people want to see women in positions of power, however it is still divided among traditional gender roles among conservatives

Everyone agrees sexual harassment is a serious issue, lowest support among men (go figure)

Women with kids do not like Trump and most people want action on gun control



Version 4: State of Economy, Support Trump, Tax Cuts, Trade, Favorite Party, Right track

Most young people say we are not on the right track across all races

Most people are happy with the performance of the economy except people whose income is <30,000-49,000

Moderate conservatives did tried to distance themselves from Trump during the election

African Americans and Latinos were typically not happy with the performance of the economy



Actual Subgroups

Socialist
Liberals

Center Left
Democrats

Fiscal
Conservativ
es/ Social
Liberals

Immigration
/ Economy
Conservativ
es

Religious
Conservativ
es

Next Steps

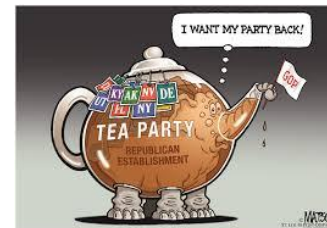
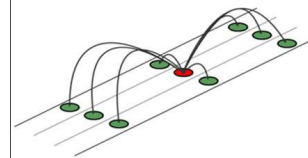
Try other surveys to confirm results

Try the autoimputer or any multi imputation technique and see if that performs better than Datawig

Find past surveys to prove this has changed over time



Imputation





Conclusions

Majority/Minority race relations are deteriorating

The divide between Republicans and Democrats has never been larger

Healthcare continues to dominate the conversation

The largest gap is between generations

America is becoming more divided

The most consistently liberal demographic are young black voters

Most people are wary of investigations

The divide between the rich and poor continues

America is hard to predict

Most Conservative: White Voters 30-59

Trump is involved in everything whether republicans like it or not

Republicans are becoming more homogenized