A Statisical Prediction of 2020 American Federal Election Result Based on Logistic Regression Model and Post-Stratification Method

Yutong Wu, Yao He, Yujie Wu, Xinxuan Lin

November 2, 2020

Model

The United States Presidential election, which happens once every four years, will be held on November 3rd, 2020. Plenty of news medias online are predicting the winner of this election and many (e.g. BBC news and CTV news) of them are saying that Biden has the lead. Here we are also interested in the result of this election and we are going to predict it by firstly created a model using the survey data, then perform a post-stratification to a census data set in order to examine our model and make the prediction. In the below sections we will explain the details of the model, as well as the details of the post-stratification process.

Model Specifics

We will be using a logistic regression model to model the proportion of voters who will vote for Donald Trump and another logistic regression model to model the proportion of voters who will vote for Joe Biden. In order to make our models to take more aspects into account, we choose gender, census region, race ethnicity, household income and age as variables to construct our model. The main reason for chosing the logistic regression is that the dependent variables that we are using are both binary variables (vote_trump and vote_biden). We want to make a prediction that to what extend can Trump/Biden win and it really makes no sense if the predicted value is greater than 1 or less than 0. Specifically, the variables we used are:

- "Gender" is a categorical variable with two possible values: "Male" and "Female". The main reason for choosing this variable is that Donald Trump has made some controversial comments about women, which may cause dissatisfaction from women.
- "Census region", which is a categorical variable that consists four possible values: "Midwest", "south", "West" and "Northeast". From the map given by CNN news, we can observe that people from different states have different preferences on the next president. However, there are more than 50 states in US and the number is way more than what we needed. One possible alternative solution is to use the census region. We can clearly see that people from west and east are more trend to vote Biden, while people from the center area are prefer to Trump.
- "Race ethnicity" is a categorical variable with possible values: "White", "Asian (Asian Indian)", "Asian (Vietnamese)", "Asian (Chinese)", "Asian (Korean)", "Asian (Japanese),"Some other race“,"Asian (Filipino)“,"Asian (Other)“,"Pacific Islander (Native Hawaiian)“,"American Indian or Alaska Native“,"Pacific Islander (Other)“,"Pacific Islander (Samoan)" and "Pacific Islander (Guamanian)". This represents the race of a given person and we are choosing this variable since Trump and Biden has totally different race policies.
- "Household income", which is a categorical data, the variables represents different intervals of household income.
- "Age" is a numerical variable that represents the age of given person.

These are the logistic regression models we are using:
For Trump:

$$log(\frac{p_{tp}}{1-p_{tp}}) = \beta_0 + \beta_1 x_{age} + \beta_2 x_{male} + \beta_3 x_{NE} + \beta_4 x_S + \beta_5 x_W + \beta_6 x_{ind} + \beta_7 x_{cn} + \beta_8 x_{fil} + \beta_9 x_{jap} + \beta_{10} x_{kor} + \beta_{11} x_{oa} + \beta_{12} x_{vit}$$

$$+\beta_{13} x_{baa} + \beta_{14} x_{gua} + \beta_{15} x_h + \beta_{16} x_{opi} + \beta_{17} x_{sa} + \beta_{18} x_{or} + \beta_{19} x_w$$

$$+\beta_{20} x_{125000} + \beta_{21} x_{15000} + \beta_{22} x_{150000} + \beta_{23} x_{175000} + \beta_{24} x_{20000} + \beta_{25} x_{200000} + \beta_{26} x_{25000} + \beta_{27} x_{250000} + \beta_{28} x_{30000} + \beta_{29} x_{35000} + \beta_{30} x_{40000}$$

$$+\beta_{31} x_{45000} + \beta_{32} x_{50000} + \beta_{33} x_{55000} + \beta_{34} x_{60000} + \beta_{35} x_{65000} + \beta_{36} x_{70000} + \beta_{37} x_{75000} + \beta_{38} x_{80000} + \beta_{39} x_{85000} + \beta_{40} x_{90000}$$

$$+\beta_{41} x_{95000} + \beta_{42} x_{14999} + \epsilon$$

For Biden:

$$log(\frac{p_{bd}}{1-p_{bd}}) = \beta_A + \beta_B x_{age} + \beta_C x_{male} + \beta_D x_{NE} + \beta_E x_S + \beta_F x_W + \beta_G x_{ind} + \beta_H x_{cn} + \beta_I x_{fil} + \beta_J x_{jap} + \beta_K x_{kor} + \beta_L x_{oa} + \beta_M x_{vit}$$

$$+\beta_N x_{baa} + \beta_O x_{gua} + \beta_P x_h + \beta_Q x_{opi} + \beta_R x_{sa} + \beta_S x_{or} + \beta_T x_w$$

$$+\beta_U x_{125000} + \beta_V x_{15000} + \beta_W x_{150000} + \beta_X x_{175000} + \beta_Y x_{20000} + \beta_Z x_{200000} + \beta_{AA} x_{25000} + \beta_{AB} x_{250000} + \beta_{AC} x_{30000} + \beta_{AD} x_{35000} + \beta_{AE} x_{40000}$$

$$+\beta_{AF} x_{45000} + \beta_{AG} x_{50000} + \beta_{AH} x_{55000} + \beta_{AI} x_{60000} + \beta_{AJ} x_{65000} + \beta_{AK} x_{70000} + \beta_{AL} x_{75000} + \beta_{AM} x_{80000} + \beta_{AN} x_{85000} + \beta_{AO} x_{90000}$$

$$+\beta_{AP} x_{95000} + \beta_{AQ} x_{14999} + \epsilon$$

Where p_{tp} represents the proportion of voters who will vote for Donald Trump and p_{bd} represents the proportion of voters who will vote for Joe Biden. Similarly, β_0 and β_A represents the intercept of the models, and is the probability of voting for Donald Trump/Joe Biden at age 0, gender is female, census_region is Midwest, race ethnicity is American Indian or Alaska Native and the household income within the range of $100,000 to $124,999. Additionally, β_1 and β_B represents the slope of the model w.r.t. to age. So, for everyone one unit increase in age, we expect a β_1 increase in the probability of voting for Donald Trump and β_B increase in the probability of voting for Joe Biden.

Similarly,

β_2 and β_C are the slopes of Male for Donald Trump's model and Joe Biden's model respectively. If β_i is negative, it means that male has negative effect on the proportion of voters. Positive β_i indicates male has positive effect.

β_3 to β_5 and β_D to β_F are the slopes of different census regions for Donald Trump's model and Joe Biden's model respectively. If β_i is negative, it means that proportion of voters is inversely proportional to the corresponding census region. Positive β_i represents that proportion of voters is proportional to corresponding census region.

β_6 to β_{19} and β_G to β_T are the slopes of different race ethnicities for Donald Trump's model and Joe Biden's model respectively If β_i is negative, it means that proportion of voters is inversely proportional to the corresponding race ethnicity Positive β_i represents that proportion of voters is proportional to corresponding race ethnicity.

$\beta_2 0$ to $\beta_4 2$ and β_u to β_{AQ} are the slopes of different household incomes for Donald Trump's model and Joe Biden's model respectively If β_i is negative, it means that proportion of voters is inversely proportional to the corresponding household income. Positive β_i represents that proportion of voters is proportional to corresponding household income.

ϵ is the error term of the models.

Note that:
x_{age} : the age of respondent.
x_{male} : respondent who was indicated as male in column of sex.
x_{NE} : vector of respondent whose census_region is Northeast.
x_S : vector of respondent whose census_region is South.
x_W : vector of respondent whose census_region is West.
x_{ind} : vector of respondent whose race_ethnicity is Asian (Asian Indian).
x_{cn} : vector of respondent whose race_ethnicity is Asian (Chinese).
x_{fil} : vector of respondent whose race_ethnicity is Asian (Filipino).
x_{jap} : vector of respondent whose race_ethnicity is Asian (Japanese).
x_{kor} : vector of respondent whose race_ethnicity is Asian (Korean).
x_{oa} : vector of respondent whose race_ethnicity is Asian (Other).
x_{vit} : vector of respondent whose race_ethnicity is Asian (Vietnamese).
x_{baa} : vector of respondent whose race_ethnicity is Black, or African American.
x_{gua} : vector of respondent whose race_ethnicity is Pacific Islander (Guamanian).
x_h : vector of respondent whose race_ethnicity is Pacific Islander (Native Hawaiian).
x_{opi} : vector of respondent whose race_ethnicity is Pacific Islander (Other).
x_{sa} : vector of respondent whose race_ethnicity is Pacific Islander (Samoan).
x_{or} : vector of respondent whose race_ethnicity is Some other race.
x_w : vector of respondent whose race_ethnicity is White.
x_{125000} : vector of respondent whose household_income is between $125,000 to $149,999.
x_{15000} : vector of respondent whose household_income is between $15,000 to $19,999.
x_{150000} : vector of respondent whose household_income is between $150,000 to $174,999.

x_{175000}: vector of respondent whose household_income is between $175,000 to $199,999.
x_{20000}: vector of respondent whose household_income is between $20,000 to $24,999.
x_{200000}: vector of respondent whose household_income is between $200,000 to $249,999.
x_{25000}: vector of respondent whose household_income is between $25,000 to $29,999.
x_{250000}: vector of respondent whose household_income is $250,000 and above.
x_{30000}: vector of respondent whose household_income is between $30,000 to $34,999.
x_{35000}: vector of respondent whose household_income is between $35,000 to $39,999.
x_{40000}: vector of respondent whose household_income is between $40,000 to $44,999.
x_{45000}: vector of respondent whose household_income is between $45,000 to $49,999.
x_{50000}: vector of respondent whose household_income is between $50,000 to $54,999.
x_{55000}: vector of respondent whose household_income is between $55,000 to $59,999.
x_{60000}: vector of respondent whose household_income is between $60,000 to $64,999.
x_{65000}: vector of respondent whose household_income is between $65,000 to $69,999.
x_{70000}: vector of respondent whose household_income is between $70,000 to $74,999.
x_{75000}: vector of respondent whose household_income is between $75,000 to $79,999.
x_{80000}: vector of respondent whose household_income is between $80,000 to $84,999.
x_{85000}: vector of respondent whose household_income is between $85,000 to $89,999.
x_{90000}: vector of respondent whose household_income is between $90,000 to $94,999.
x_{95000}: vector of respondent whose household_income is between $95,000 to $99,999.
x_{14999}: vector of respondent whose household_income is Less than $14,999.

vote_trump

term	estimate	std.error	statistic	p.value
(Intercept)	-0.7466521	0.2708746	-2.7564489	0.0058433
age	0.0105090	0.0018102	5.8055245	0.0000000
genderMale	0.4240564	0.0581318	7.2947342	0.0000000
census_regionNortheast	-0.1624046	0.0898653	-1.8071991	0.0707312
census_regionSouth	0.3126448	0.0777345	4.0219566	0.0000577
census_regionWest	-0.1295545	0.0870790	-1.4877817	0.1368085
race_ethnicityAsian (Asian Indian)	-0.5326172	0.3406740	-1.5634222	0.1179533
race_ethnicityAsian (Chinese)	-1.2157582	0.4034368	-3.0135036	0.0025825
race_ethnicityAsian (Filipino)	-0.0239774	0.4039740	-0.0593537	0.9526704
race_ethnicityAsian (Japanese)	-1.0119644	0.6227269	-1.6250535	0.1041512
race_ethnicityAsian (Korean)	-0.2391505	0.6917096	-0.3457383	0.7295394
race_ethnicityAsian (Other)	-0.7770044	0.4953243	-1.5686780	0.1167230
race_ethnicityAsian (Vietnamese)	-0.7328805	0.6813795	-1.0755834	0.2821136
race_ethnicityBlack, or African American	-1.9494737	0.2695057	-7.2335164	0.0000000
race_ethnicityPacific Islander (Guamanian)	13.2014898	535.4112566	0.0246567	0.9803288
race_ethnicityPacific Islander (Native Hawaiian)	-0.2642703	0.7402971	-0.3569788	0.7211077
race_ethnicityPacific Islander (Other)	-13.2833875	197.4332472	-0.0672804	0.9463585
race_ethnicityPacific Islander (Samoan)	-13.4246119	374.9228947	-0.0358063	0.9714368
race_ethnicitySome other race	-0.6332962	0.2634676	-2.4036968	0.0162302
race_ethnicityWhite	0.0655850	0.2384372	0.2750620	0.7832686
household_income$125,000 to $149,999	-0.0480498	0.1525988	-0.3148769	0.7528551
household_income$15,000 to $19,999	-0.4766550	0.1614740	-2.9518992	0.0031583

household_income$150,000 to $174,999	-0.1627126	0.1842384	-0.8831635	0.3771479
household_income$175,000 to $199,999	0.3140245	0.2219912	1.4145806	0.1571915
household_income$20,000 to $24,999	-0.0911866	0.1570640	-0.5805695	0.5615306
household_income$200,000 to $249,999	0.6223445	0.1997762	3.1152075	0.0018382
household_income$25,000 to $29,999	-0.2986818	0.1570213	-1.9021731	0.0571485
household_income$250,000 and above	0.2310988	0.2116413	1.0919362	0.2748611
household_income$30,000 to $34,999	-0.3137536	0.1562651	-2.0078288	0.0446615
household_income$35,000 to $39,999	-0.2906401	0.1628695	-1.7844974	0.0743429
household_income$40,000 to $44,999	-0.3702276	0.1708093	-2.1674907	0.0301975
household_income$45,000 to $49,999	-0.1831347	0.1636010	-1.1193984	0.2629702
household_income$50,000 to $54,999	-0.1628805	0.1550688	-1.0503756	0.2935454
household_income$55,000 to $59,999	-0.1485219	0.1936838	-0.7668265	0.4431847
household_income$60,000 to $64,999	-0.4004036	0.1923218	-2.0819462	0.0373474
household_income$65,000 to $69,999	-0.2797773	0.2117494	-1.3212664	0.1864126
household_income$70,000 to $74,999	-0.1872242	0.1864339	-1.0042391	0.3152634
household_income$75,000 to $79,999	-0.0528074	0.1849966	-0.2854507	0.7752989
household_income$80,000 to $84,999	-0.4854948	0.2277054	-2.1321187	0.0329971
household_income$85,000 to $89,999	-0.2944993	0.2455014	-1.1995829	0.2303014
household_income$90,000 to $94,999	-0.0888777	0.2620992	-0.3390995	0.7345347
household_income$95,000 to $99,999	-0.4888970	0.2007216	-2.4356970	0.0148631
household_incomeLess than $14,999	-0.3980824	0.1270400	-3.1335211	0.0017272

vote_biden

term	estimate	std.error	statistic	p.value
(Intercept)	-0.5493469	0.2764078	-1.9874507	0.0468725
age	-0.0020619	0.0017557	-1.1744191	0.2402272
genderMale	-0.3151989	0.0564554	-5.5831447	0.0000000
census_regionNortheast	0.0735511	0.0869895	0.8455160	0.3978228
census_regionSouth	-0.2511399	0.0759724	-3.3056723	0.0009475
census_regionWest	0.0159866	0.0842179	0.1898239	0.8494471
race_ethnicityAsian (Asian Indian)	0.7737619	0.3279838	2.3591470	0.0183170
race_ethnicityAsian (Chinese)	1.1407700	0.3573614	3.1922023	0.0014119
race_ethnicityAsian (Filipino)	0.4394549	0.4043596	1.0867924	0.2771286
race_ethnicityAsian (Japanese)	1.6108397	0.5848990	2.7540477	0.0058863
race_ethnicityAsian (Korean)	1.1359340	0.6859421	1.6560202	0.0977177
race_ethnicityAsian (Other)	0.6464691	0.4364823	1.4810887	0.1385829
race_ethnicityAsian (Vietnamese)	0.4332452	0.6766468	0.6402826	0.5219889
race_ethnicityBlack, or African American	1.7632058	0.2585779	6.8188563	0.0000000

race_ethnicityPacific Islander (Guamanian)	-10.7280656	196.9679261	-0.0544661	0.9565639
race_ethnicityPacific Islander (Native Hawaiian)	0.9701684	0.6954510	1.3950205	0.1630097
race_ethnicityPacific Islander (Other)	2.6040920	1.1123508	2.3410709	0.0192285
race_ethnicityPacific Islander (Samoan)	0.7237738	1.4519638	0.4984793	0.6181463
race_ethnicitySome other race	0.7446719	0.2638762	2.8220498	0.0047718
race_ethnicityWhite	0.2596597	0.2455554	1.0574383	0.2903116
household_income$125,000 to $149,999	0.1980848	0.1527467	1.2968190	0.1946935
household_income$15,000 to $19,999	0.1996743	0.1561857	1.2784414	0.2010938
household_income$150,000 to $174,999	0.2894369	0.1836175	1.5763033	0.1149559
household_income$175,000 to $199,999	-0.1291920	0.2278120	-0.5670991	0.5706469
household_income$20,000 to $24,999	0.1710727	0.1549262	1.1042212	0.2694972
household_income$200,000 to $249,999	-0.4346854	0.2094585	-2.0752817	0.0379604
household_income$25,000 to $29,999	0.1706654	0.1543046	1.1060289	0.2687140
household_income$250,000 and above	0.0738400	0.2111191	0.3497553	0.7265224
household_income$30,000 to $34,999	0.1913584	0.1540472	1.2422062	0.2141605
household_income$35,000 to $39,999	0.2622072	0.1591714	1.6473261	0.0994910
household_income$40,000 to $44,999	0.3235097	0.1673930	1.9326355	0.0532811
household_income$45,000 to $49,999	0.1595545	0.1627273	0.9805024	0.3268382
household_income$50,000 to $54,999	0.2451219	0.1529219	1.6029220	0.1089519
household_income$55,000 to $59,999	0.1476896	0.1908254	0.7739518	0.4389593
household_income$60,000 to $64,999	0.2912115	0.1875564	1.5526609	0.1205042
household_income$65,000 to $69,999	0.2604118	0.2091706	1.2449733	0.2131416
household_income$70,000 to $74,999	0.3712785	0.1828800	2.0301753	0.0423387
household_income$75,000 to $79,999	0.0751593	0.1842516	0.4079165	0.6833350
household_income$80,000 to $84,999	0.4962945	0.2194876	2.2611510	0.0237499
household_income$85,000 to $89,999	0.6030675	0.2402547	2.5101176	0.0120691
household_income$90,000 to $94,999	0.2837044	0.2629773	1.0788173	0.2806692
household_income$95,000 to $99,999	0.5428345	0.1950651	2.7828381	0.0053886
household_incomeLess than $14,999	0.0669351	0.1241848	0.5389962	0.5898894

Post-Stratification

In the above section, we created two models to predict the proportion of voters who will vote Donald Trump and Joe Biden separately. In this section, we will perform post-stratification analyses, which “incorporating population distributions of variables into survey estimates” (R. J. A. Little) for both Trump’s model and Biden’s model to estimate the proportion of voters who will vote for them. In the analyses, we will group similar unit together in order to reduce the variance. The variables we are using to construct the groups are: age, gender, household_income, race_ethnicity and census_region.

- “Gender”: The main reason for choosing this variable is that Donald Trump has made some controversial comments about women, which may cause dissatisfaction from women.

- “Census region”: From the map given by CNN news, we can observe that people from different states have different preferences on the next president. However, there are more than 50 states in US and the number is way more than what we needed. One possible alternative solution is to use the census region. We can clearly see that people from west and east are more trend to vote Biden, while people from the center area are prefer to Trump.
- “Race ethnicity”: We are choosing this variable since Trump and Biden has totally different race policies.
- “Household income”: Donald Trump and Joe Biden has different plans towards American economics. For example, their tax plans are different. The effects of these plans on people from different income level are different and people from different income level may prefer different ones. Thus, we decide to include “Household income” as a factor to split the data.
- “Age”: We are choosing this variable since we think younger people pay more attention to immediate benefits while older people usually make long-term plans.

We will then weight each proportion estimate (within each bin) by the respective population size of that bin and sum those values and divide that by the entire population size.

```
## # A tibble: 1 x 1
##   alp_predict_for_Trump
##                   <dbl>
## 1                 0.370
```

```
## # A tibble: 1 x 1
##   alp_predict_for_Biden
##                   <dbl>
## 1                 0.487
```

Results

Based on Table 1, the slope of age is 0.01051 which shows that as age increases, people are more likely to vote for Trump. Moreover, the slope of male is 0.42406 which suggests that men are more likely to vote for Trump. Similarly, the slope of South census region, White and Guamanian race ethnicities, household income from $175,000 to $199,999, household income between $200,000 and $249,999, and household income above $250,000 are all positive which indicate that South region people, White and Guamanian people, and people who has income above $175,000 are more likely to vote for Trump.

Table 2 shows that the slopes for age and male are both negative, likely showing that with the increase of age, most people are in favour of not voting for Biden; this includes most males as well. Moreover, different from Figure 1, except slopes for South region, race of Guamanian and White, and slopes for household income above $175,000, other slopes of corresponding variable are all positive, which indicate that people who have a household income below $175,000 are more liable to vote Biden; people who lived outside of the South region, and people whose race are not White or Guamanian are more likely to vote Biden.

We applied the post-stratification method to determine the proportion of voters who prefer voting for the Republican Party, and the proportion of voters who are likely to vote for the Democratic Party. It is modelled by the logistic model and is using the respondents’ gender, age, household income, census region, and race ethnicity as predictors. According to this analysis, we estimate that the proportion of voters who are inclined to vote for the Republican Party is around 0.370 while the proportion of voters who favoured the Democratic Party is 0.487.

Discussion

In this study, we use the survey data from Nationscape Data Set to fit two logistic regression models. One for the probability of voting for Trump and the other for the probability of voting for Biden. Then we use the census data from IPUMS data to predict the US election results for both Trump and Biden. The prediction variables that we used include age, gender, household income, race/ethnicity and region. The results show that males, Southern region people, White and Guamanian people as well as rich people (people whose income are higher than $175,000) are more likely to vote for Trump, while majority of the the rest of the US citizens are more likely to vote for Biden. This actually makes sense according to the previous news and Biden and Trump’s election campaign. Trump is a controversial character in both gender equality and race equality since he has some improper words towards these topics. It is not doubtful that females and non-white and non-guamanian people are not in favour of Trump. Meanwhile, Trump is dedicated to cut corporate taxes, which indeed is in the interest of the rich people because the rich tend to enjoy more benefit from the tax cut. In addition to that, Trump is more liked in the south region because Trump’s standpoints are like the ones of the southern people. Based on the above reasons, Trump’s supporters tend to be males, Southern region people, White and Guamanian people as well as rich people. On the other hand, Biden stands on the opposite side of Trump, and therefore tends to gain votes from citizens other than males, Southern region people, White and Guamanian people as well as rich people.

From the model that we fit and the variables that we think may affect the election result, we see that the probability of the public voting for Trump is 0.370 and the probability for Biden is 0.487. With that being said, based on the factors of xxx, the election result might be that Biden wins the election this year, with an about 10% lead.

Weaknesses & Next Steps

Although we were able to draw conclusions from the data, there is still drawback in our dataset that reduces the accuracy of our models. The race & ethnicity category in our survey data and census data does not perfectly match. Racial categories are not backed by science and therefore it is very hard to categorize people in a way that is inclusive and accurate. Some identify their races biologically while some may identify their races based on their born city. In our case, for example, to match the “two major races” and “three or more major races” in census data with the survey data, we class these races as “Some other race”. It is not a scientifically accurate categorization but such bias can not be avoided. Additionally, a significant variable which will influence our model in our data set is that around 10% of the respondents selected “not sure” when being asked about their selection of president. For a further survey, it is essential to compare the estimate results with the actual election results to improve our model for future election. To complete this, we will use the fresh data of the 2020 election result to repeat the process in the model part and compare the outcome. Notice at that point, there will be no individuals who are uncertain about their choice on US presidential election. It would be interesting to find out how and to what extent the uncertainty affect our models.

Appendix: Github

https://github.com/wuyujie1/STA304PS3 (https://github.com/wuyujie1/STA304PS3)

References

2020 United States presidential election. (2020, November 02). Retrieved November 01, 2020, from https://en.wikipedia.org/wiki/2020_United_States_presidential_election (https://en.wikipedia.org/wiki/2020_United_States_presidential_election)

Slaughter, G. (2020, October 29). Five days until U.S. election: Polls show Biden leading, but Trump still sees path to victory. Retrieved November 01, 2020, from https://www.ctvnews.ca/world/america-votes/five-days-until-u-s-election-polls-show-biden-leading-but-trump-still-sees-path-to-victory-1.5164757 (https://www.ctvnews.ca/world/america-votes/five-days-until-u-s-election-polls-show-biden-leading-but-trump-still-sees-path-to-victory-1.5164757)

The Visual and Data Journalism Team. (2020, November 01). US election 2020 polls: Who is ahead - Trump or Biden? Retrieved November 01, 2020, from https://www.bbc.com/news/election-us-2020-53657174 (https://www.bbc.com/news/election-us-2020-53657174)

Prasad, R. (2019, November 29). How Trump talks about women - and does it matter? Retrieved November 01, 2020, from https://www.bbc.com/news/world-us-canada-50563106 (https://www.bbc.com/news/world-us-canada-50563106)

The Road to 270: Interactive Electoral College maps. (n.d.). Retrieved November 01, 2020, from https://www.cnn.com/election/2020/electoral-college-interactive-maps (https://www.cnn.com/election/2020/electoral-college-interactive-maps)

Trump and Biden couldn’t be more different on the complicated issue of race. (2020, August 18). Retrieved November 01, 2020, from https://www.latimes.com/politics/story/2020-08-06/trump-biden-race-policy (https://www.latimes.com/politics/story/2020-08-06/trump-biden-race-policy)

Post-Stratification. (n.d.). Retrieved November 01, 2020, from https://methods.sagepub.com/reference/encyclopedia-of-survey-research-methods/n388.xml (https://methods.sagepub.com/reference/encyclopedia-of-survey-research-methods/n388.xml)

Little, R. J. (1993). Post-Stratification: A Modeler’s Perspective [Abstract]. Journal of the American Statistical Association, 88(423), 1001-1012. doi:10.1080/01621459.1993.10476368 (doi:10.1080/01621459.1993.10476368)

Gandel, S. (2020, October 30). Comparing the Biden and Trump tax plans: Will you pay more? Retrieved November 01, 2020, from https://www.cbsnews.com/news/biden-tax-plan-comparison-trump/ (https://www.cbsnews.com/news/biden-tax-plan-comparison-trump/)

Steven Ruggles, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas and Matthew Sobek. IPUMS USA: Version 10.0 ACS 2018. Minneapolis, MN: IPUMS, 2020. https://doi.org/10.18128/D010.V10.0 (https://doi.org/10.18128/D010.V10.0)

Tausanovitch, Chris and Lynn Vavreck. 2020. Democracy Fund + UCLA Nationscape, October 10-17, 2019 (version 20200625). Retrieved from https://www.voterstudygroup.org/publication/nationscape-data-set (https://www.voterstudygroup.org/publication/nationscape-data-set)

U.S. Department of Commerce Economics and Statistics Administration U.S. Census Bureau. (n/a). Census Regions and Divisions of the United States. Retrieved November 2, 2020, from https://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf (https://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf)

Jenée Desmond-Harris, J., & Caswell, E. (2015). The myth of race, debunked in 3 minutes. Vox. https://www.vox.com/2015/1/13/7536655/race-myth-debunked (https://www.vox.com/2015/1/13/7536655/race-myth-debunked)