Avi Gupta
agupta07@stanford.edu
November 14, 2020

Understanding Voting Trends in Georgia to Combat Suppression

Introduction

On Election "Day" 2020, neither of Georgia's incumbent senators drew a simple majority of votes, sending both of their races into special runoff elections in January 2021 that will determine party control of the Senate. Incumbent Republican Senators David Perdue and Kelly Loffler are being respectively challenged by Democrats John Ossof and Rev. Dr. Raphael Warnock.

Senate control will be essential for President-Elect Joe Biden to freely carry out his policy agenda and push through Supreme Court nominations as he sees fit. According to the New York Times, "tens of millions of dollars in campaign cash are expected to pour into the state to fund a marathon of political advertising."

Given that national attention is turning towards Georgia, I decided to work on understanding voting data from the state. I became particularly interested in understanding voter turnout in terms of racial demographics after hearing about the recent work of Stacey Abrams, who is being widely credited with transforming Georgia from a historically Republican state into a political battleground state. After narrowly losing her bid for Georgia's governor seat in an election tainted by stains of voter suppression targeting Black voters, Abrams launched Fair Fight, an organization that encourages voter participation and fights targeted voter suppression. Fair Fight has registered 800,000 first-time voters over the past two years.

My goal with this project was to first create visuals that might help volunteers at groups like Fair Fight quickly understand how voting trends are shifting in Georgia on a county by county basis and for races that may be systematically suppressed from exercising their right to vote. I additionally wanted to understand if changes in voter turnout for 4 race groups from 2018 to 2020 were statistically significant in order to more tangibly grasp where the attention of voting rights organizations could be focused.

Finding and Processing Data

I uncovered election data on the website of the Georgia Secretary of State. To illustrate change in voter behavior in a relatively short period of time and lightly speculate about January 2021, I felt it was appropriate to look at the change from 2018 to 2020 rather than 2016 to 2020. Thus, I sourced two datasets: general election turnout by demographics and county in November 2018, and November 2020.

Both 2016 and 2018 voter datasets I found provided voter counts by race, gender, and county. I processed the data within Excel to only account for race and county by summing these vote counts across all provided genders: male, female, and unknown. For example, I summed the male, female, and other black voters in Appling County to create a single data point representing the number of black voters from Appling County.

The four race and ethnicity categories I considered were: White, Black, Hispanic/Latinx and Other. I created "Other" by combining data from American Indian, Alaskan Native, Asian, Pacific Islander, and Other (this second "Other" has no further breakdown) voters.

Visualizing Voting Trends

I used this data along with NumPy and Matplotlib in Python to plot four stacked bar graphs -- one for each race -- to illustrate the percent of each county's population that voted in 2018 and 2020.

In each graph, counties are ordered alphabetically but labeled numerically from 1 to 155. County #1 is Appling County, and County #155 is Worth County. Georgia consists of 159 counties, but I omitted Wheeler, Dawson, Mitchell, and Gilmer counties because the number of voters in the 2020 election exceeded the projected population for 2020, which led to a percentage greater than 1. Figures are displayed in the Appendix.

These visualizations could be useful in many ways. At a high level, it becomes clear from a quick glance that the probability of a white person voting in Georgia is generally higher than the probability of a person from a minority demographic voting. At a lower level, the county by county visual quickly differentiates certain counties for having relatively high or low voter turnout within their demographic. Turning our attention to Figure 2, for example, County #50's (Echols County) Black voter turnout skyrockets above the norm. One could research Echols County to learn what Echols is doing right and apply their strategy across the state. In a similar manner, one could look at Figure 3 and notice that County #126 (Stewart County) has a 1% Hispanic/Latinx voter turnout. They might then investigate why this is the case and focus their attention on alleviating any potential suppressive tactics or burdens that are in place for these voters.

Statistical Significance of Turnout Changes

Another important question that I set out to investigate was whether the observed change in the average voting probability of each race from 2018 to 2020 was statistically significant. I did not use np.mean() to calculate the average. Instead, I calculated a weighted average for that race group across all counties, weighted by the proportion of Georgia's population represented by that county.

I assumed a null hypothesis that the two voting samples were drawn from the same year, and used bootstrapping to estimate a p-value for the difference in the weighted averages of voter turnout probabilities between 2018 and 2020 across all counties.

My program output was:
Observed Difference in Weighted Avg. for White: 0.26992597065508567
Bootstrapped p-value for White: **0.0**
Observed Difference in Weighted Avg. for Black: 0.049336344870597515
Bootstrapped p-value for Black: **0.2877**
Observed Difference in Weighted Avg. for Hispanic: 0.15785979539440761
Bootstrapped p-value for Hispanic: **0.0001**
Observed Difference in Weighted Avg. for Other: 0.18271565880875718
Bootstrapped p-value for Other: **0.0246**

For the White, Hispanic/Latinx, and Other, we can reject the null hypothesis. The difference in average voting probability between 2018 and 2020 was statistically significant at the 0.05 level. However, the same cannot be said for Black Georgians. We learn from this finding that much work still remains to support and rally the Black vote. Organizations such as Fair Fight and Black Votes Matter are crucial in this fight to alleviate voter suppression in its many forms.

In Figure 5, I also created a bar graph for the weighted average of voter turnout percentages by race, with a side-by-side comparison of the two counties. For this graph, I added error bars using the standard deviation of the percentages that were averaged to calculate the value of each bar. I hope this graph helps provide a visual backing to the p-value analysis above, as the difference between the two years for Black voters looks noticeably smaller than that of other groups.

Appendix

Figure 1: White Voter Turnout by County in 2018 and 2020

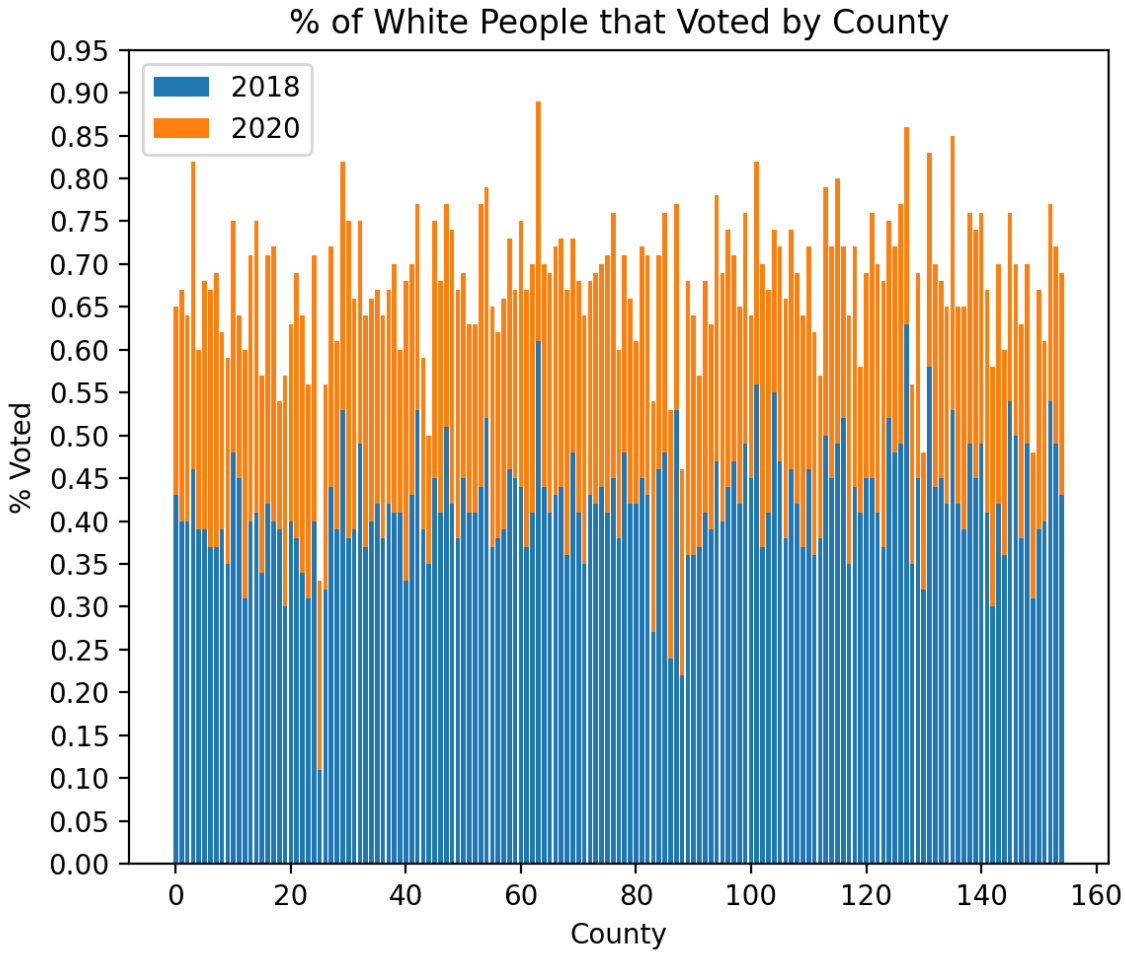

Figure 2: Black Voter Turnout by County in 2018 and 2020

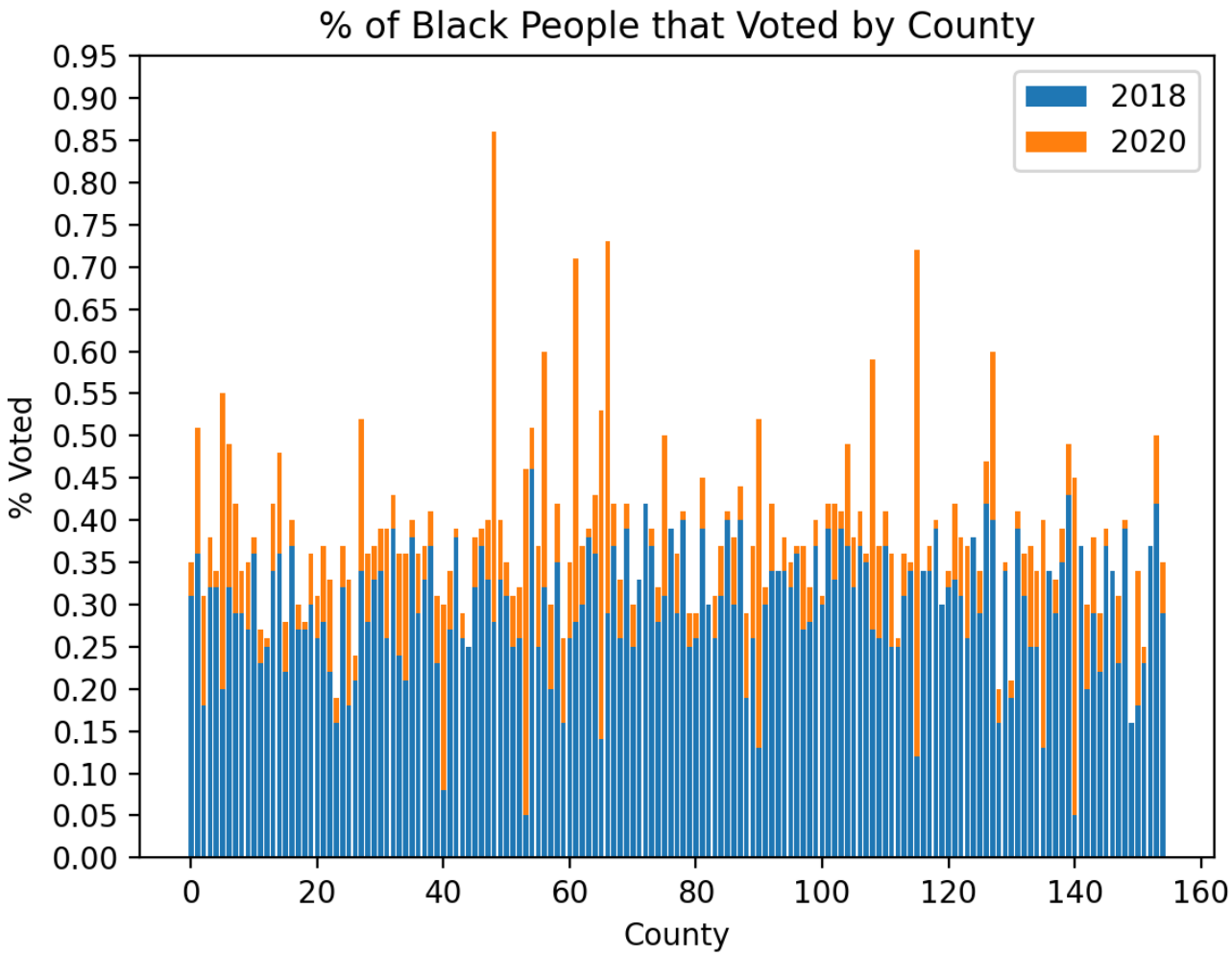

Figure 3: Hispanic Voter Turnout by County in 2018 and 2020

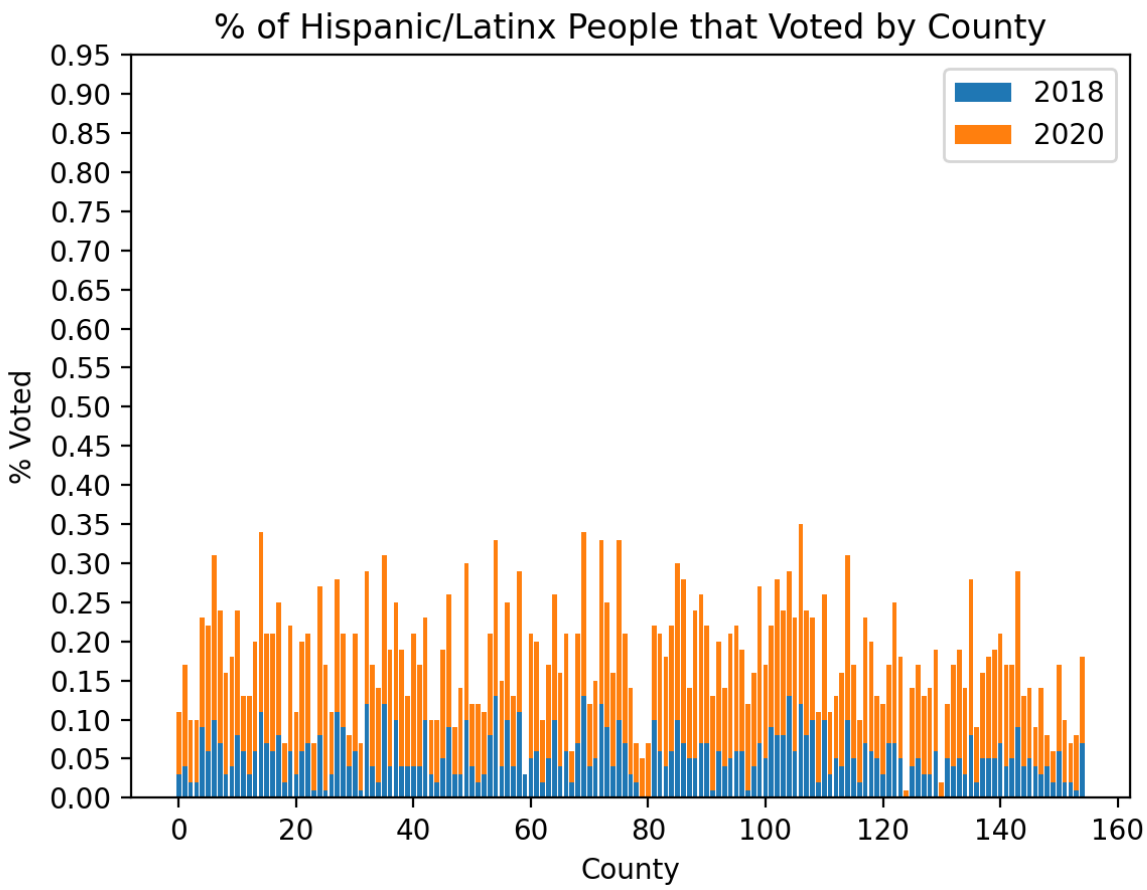

Figure 4: Other Voter Turnout by County in 2018 and 2020:

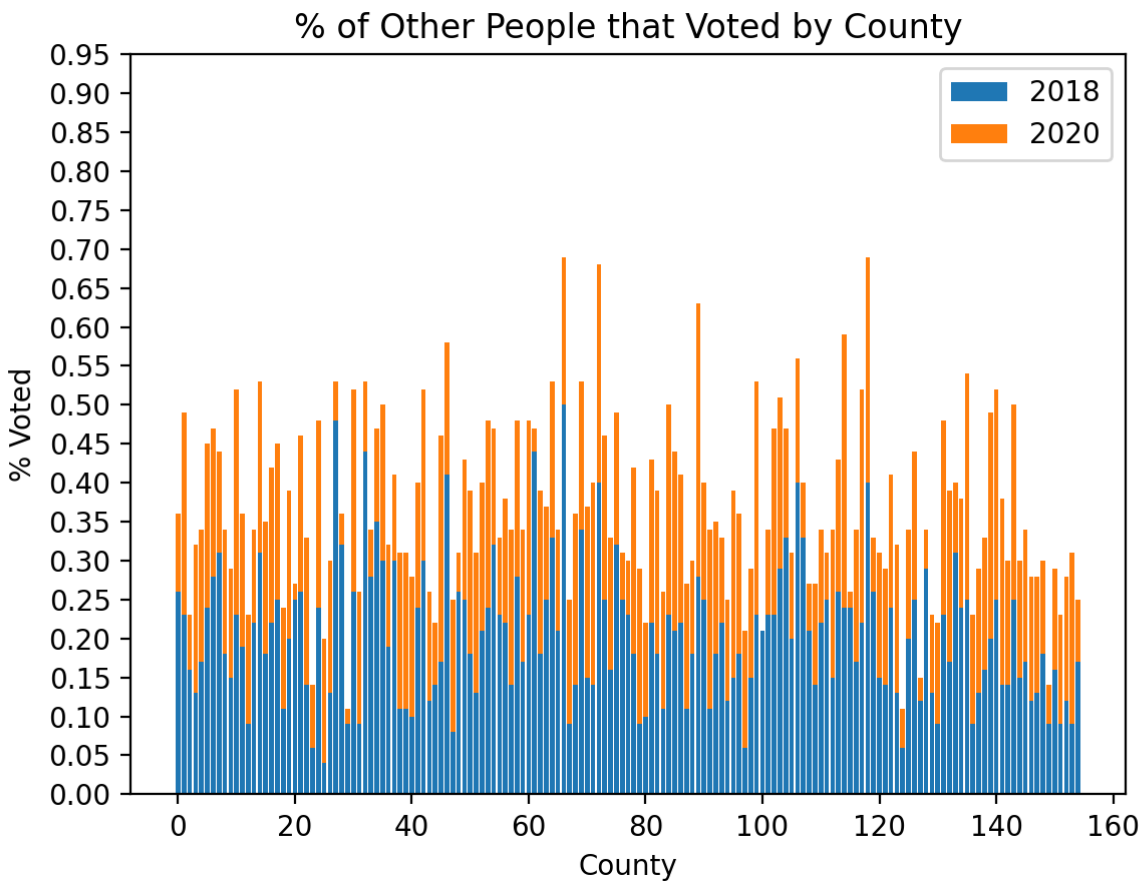

Figure 5: (Weighted) Average Voter Turnout By Race