Avocado Toast and the Aged

Lab 1, Question 1: Are Democratic voters older or younger than Republican voters in 2020?

Hannah Gross, Matt Kinkley, Jacquie Nesbitt, Huyette Spring

Contents

1 Importance and Context

In recent elections, Democrats have targeted younger voters while Republicans have targeted older voters. This strategy is in line with the popular narrative that the Democratic Party is the party of youth. We have seen several successful campaigns led by Democrat candidates to mobilize young voters. A notable example is Jon Ossoff's successful use of TikTok to motivate young voters in the contentious Georgia run-off campaign in early 2021 (Posner). But, are Democratic voters actually younger than Republican voters?

We sought to investigate these claims by asking: Are Democratic voters older or younger than Republican voters in the 2020 Election? This investigation serves as a basis for future research, and may also provide invaluable information to help guide the future policy and campaign strategy for both parties.

2 Description of Data

We addressed this research question using data from the 2020 American National Election Studies (ANES) Time Series Study. This study is an observational dataset based on a combined sample of respondents who participated in the prior ANES 2016 Time Series Study and a freshly drawn cross-section of randomized addresses of eligible voters at the time of recruitment. The study is a preliminary release of combined pre and post-election data and, therefore, subject to future changes in cleaning, processing, documentation, data, and variables.

For our grouping variable, we determined whether a participant was a Democrat or a Republican using the following survey question (V201018): "What political party are you registered with, if any?" Respondents answered by selecting one out of the following four options: "Democratic party", "Republican party", "None or 'Independent' ", or "Other". This question appeared only if the respondent was registered to vote or registered to vote in a state where the party may be registered at the time of surveying. We believe this is an appropriate and conservative variable to operationalize the research question because we wanted to focus on individuals who have self-declared their party and not their ideological leanings.

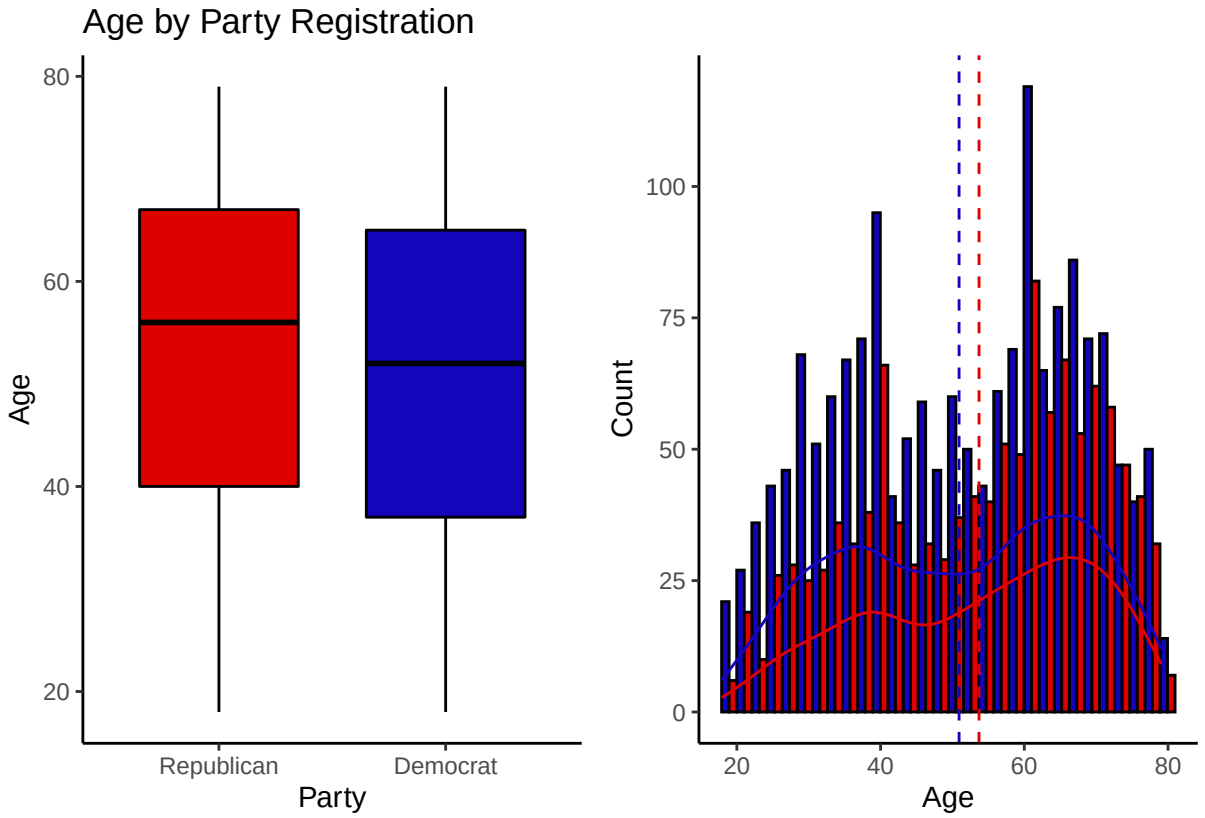
We considered alternative party identification variables, but we decided against using a variable associated with party identification because of the iterative process of asking study participants to define their partisan leanings. The iterative process compels respondents to identify with one of the two major US political parties. For this research question, we did not want to use these potentially inaccurate responses. Rather we aimed to identify Republicans and Democrats by how they self-identified in their voter registration, a process independent of the ANES survey.

Another potential approach to operationalizing party affiliation was to create a composite variable that would include the sum of votes across all political races. However, this would have required more time and resources available but is worthy of future exploration.

For our response variable, we used the age summary variable (V201507x). The survey asked for the respondent's date of birth which is calculated into the age summary variable (V201507x). We felt that this was the most appropriate variable because it gave us the exact age of a respondent and we could not identify any other compelling alternatives.

We operationalized the party of registration variable by removing all empty values and dropping all answers not equal to the Democratic Party or Republican Party. As for the age variable, we removed all empty values and non-applicable values. The ANES dataset de-identified ages 80 and above and masked them all as 80 years old. After further investigation, we found the number of respondents in their 80's or above to be 401 observations or about 4.8% of the sample and were roughly equally divided between Democrats and Republicans so we decided to drop them. Thus, after subsetting the data for Democrats and Republicans between 18 and 79 years old, we had 2,869 observations.

The figure below depicts the distribution of respondent's age by party registration. According to these graphs, registered Democrat respondents are skewed slightly younger than registered Republican respondents in 2020.



3 Most appropriate test

First, we deduced a parametric test was appropriate because the age variable is on a metric scale. Since we were only looking at age, we knew that the data was unpaired. Therefore, we decided on a Welch's t-test, implemented in R using the default t.test. Also, we decided to adopt a more conservative statistical approach by conducting a two-tailed t-test. A one-tailed t-test limits the scope of the test and may hide important statistical results.

The Welch's t-test requires the following three assumptions to be true.

1. The data is a random (I.I.D.) sample: Welch's t-test requires that the data is generated through an I.I.D. sampling process. This survey is a combined sample of respondents who participated in the ANES 2016 Time Series Study and a freshly drawn cross-section of randomized addresses across 50 states and the District of Columbia of eligible voters at the time of recruitment. For the fresh cross-section, respondents were sent a series of letters to recruit one household member to complete a survey for each randomly selected address. However, it is important to note that ANES investigators detected instances where the person who took the initial screener was not the same person who completed the ANES survey. ANES investigators flagged those instances and are expected to provide updated data in the final release. This dataset meets the I.I.D. requirement because each data entry in the set does not provide any information about any other entry in the set. However, the mix-matching of individuals is a limitation in the quality of the data and statistical methods used to analyze it.

2. The scale of measurement is metric: Welch's t-test requires that the variable being measured is metric. In this analysis, we measured the age variable which fits this requirement because its values are on a metric scale. This requirement does not apply to our grouping variable, i.e. party registration.

3. Normally distributed data: Welch's t-test requires that the distribution must not be too unnormal

considering the sample size. The figure above indicates a slightly bimodal distribution (Batman!) for both parties, but with 2,869 observations, we are confident that there will be no major deviations from a normal curve due to the Central Limit Theorem. Lastly, a Welch's t-test does not require equal variances, which is why we chose it over a classical t-test.

4 Test, Results, and Interpretation

Null Hypothesis: $\mu_D = \mu_R$

Alternative Hypothesis: $\mu_D \neq \mu_R$ (two-tailed)

Rejection Criteria: $P - value \leq .05$

The null hypothesis for this test is that Democrats and Republicans have the same expected age. The alternative hypothesis is that Democrats and Republicans do not have the same expected age. We also set our rejection criteria to be less than or equal to .05.

We rejected the null hypothesis because the Welch's t-test produced a highly significant p-value of 6.21e-06, which was much less than our rejection criteria (.05). The 95% confidence interval indicates there was roughly an expected age difference between 1.57 and 3.98 years.

The test suggests the average age of Democrats is 50.9 years old and the average age of Republicans is 53.7 years, indicating that Democrats are slightly younger than Republicans, by only a few years. While we have a statistically significant result, Cohen's D is -0.17 standard deviations, indicating a small effect size. However, the distribution of ages in the sample is particularly compelling. Republican ages skew towards early 60's while Democrat ages have two peaks between 30 - 40 years and 60 - 70. This indicates that Democratic voters are more youthful than Republican voters.

5 Test Limitations

While we believe that studying age differences between the core voter bases of Republicans and Democrats could yield invaluable information to help guide the future policy and campaign strategy of both parties, we view this analysis as only a small part of a much greater study. The limitations of our test range from our approach to the data as well as the research question itself. Our decisions to limit our sample to only registered Democrats and Republicans and also discount 80+ year olds due to its encoding, were purposefully conservative choices. While these choices allowed us to avoid potentially misrepresenting survey respondents, these conservative decisions may have impacted the statistical power of our result by substantially diminishing our sample size. Furthermore, the broad scope of our research question required a national-level analysis, which may have concealed some more impactful insights. For instance, comparing the ages of Democrats and Republicans for each US state may yield a more precise and meaningful result. Despite the limitations of our test, we were able to find a statistically significant difference in the expected age of Democrats and Republicans. We hope that this finding will be impactful enough to inspire additional investigation.

6 Bibliography

American National Election Studies. 2021. ANES 2020 Time Series Study Preliminary Release: Combined Pre-Election and Post-Election Data [dataset and documentation]. March 24, 2021 version. www.electionstudies.org

Brookings Institute, "The demographic blowback that elected Donald Trump"

Center for Information & Research on Civic Learning and Engagement (CIRCLE), "Election Week 2020: Young People Increase Turnout, Lead Biden to Victory"

Kaplan, "Georgia's new Senator-elect Jon Ossoff is a millennial success story"

Kral, "Republicans Must Attract More Young Voters"

McGraw, “The GOP Is Dying Off. Literally.”

NPR - All Things Considered,“Young Voter Engagement In This Year’s Election”

Opinsky, “TIKTOK: ENGAGE YOUNG VOTERS & DELIVER THE YOUTH VOTE ON SOCIAL MEDIA”

Petrocik, “Measuring party support: Leaners are not independents”

Pew Research Center “Voters Rarely Switch Parties, but Recent Shifts Further Educational, Racial Divergence”

Posner, “Democrats Take Campaign to TikTok”