

Lab 1

Courtney Chen, Chris Rezny, Jordan Andersen

March 4, 2025

Importance and Context

The political climate has remained deeply polarized over the past three election cycles, with Democrats and Republicans seemingly holding opposing stances on nearly every major issue. One such issue that has frequently been at the center of these conflicts is education. With constant exposure to media headlines, we are often inclined to wonder whether the average voters truly hold such opposing viewpoints, or if in fact the differences are not as extreme as they are presented. We will focus specifically on analyzing whether polarization is present in party views on schooling by answering: “Do Democratic and Republican voters have different views of primary and secondary schooling?”

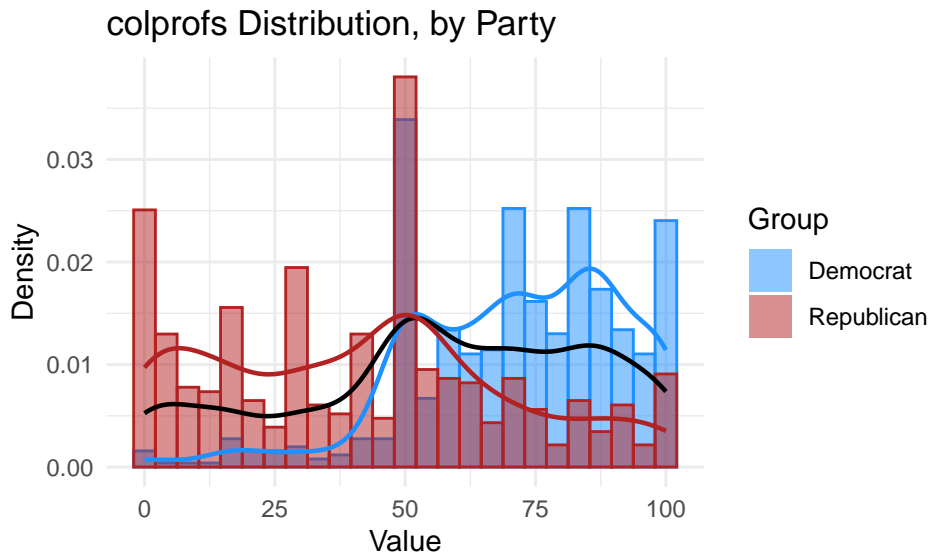
Our data has been pulled from the American National Election Studies 2024 Pilot Study, containing 1,500 registered voters in the United States. We have defined a voter as those who responded positively when asked if they did or did not vote in the most recent election. If this question was skipped, we took data from the following question which asked if the participant would show up as “voted” or “didn’t vote” if it were to be looked up in the voter registration records (`voteturn_saveface`, `voteturn_lookup`) and included positive responses. Since our goal revolved around comparing Democratic and Republican voters, great care was taken to identify voters within those parties. In alignment with literature Arceneaux & Nickerson (2009), we determined that those who identify as “Independents” but lean one way or the other should be considered a part of that particular party. To further determine party membership, we looked at the free text entry of party affiliation, and, where applicable, aligned the responses of those participants with their respective party. We also filtered out respondents who were not asked / did not respond to questions relating to their feelings towards college professors and the importance of content taught in school since the answers to these questions were our key points for analysis. We were left with the following Democratic to Republican number of respondents for the college and importance questions, respectively: 460/425, 626/565.

Data and Methodology

The data for this study is sourced from the American National Election Study (ANES), a research initiative dedicated to advancing political science through the collection of time-series voting data. Specifically, this study focuses on the ANES 2024 Pilot Study, a dataset comprising survey responses to a questionnaire that captures public opinion during the early stages of the 2024 election cycle. The questionnaire explores a broad range of political topics, including student loans, foreign policy, and more. Our analysis will center on key questions related to the education system, examining how different political parties perceive college professors and the importance of primary and secondary schooling. See below for key terms:

- Primary schooling: Grades 1-12 (first grade until last year of high school, inclusive)
- Secondary schooling: Higher education (first year of college and beyond)
- `imp_schteach`: Importance of what’s being taught in public schools (**Likert scale**: *Not at all important, Slightly important, Moderately important, Very important, Extremely important*)
- `group_colprofs`: Rating of college professors using a **feeling thermometer (0-100 scale, where 100 represents the most favorable view)**

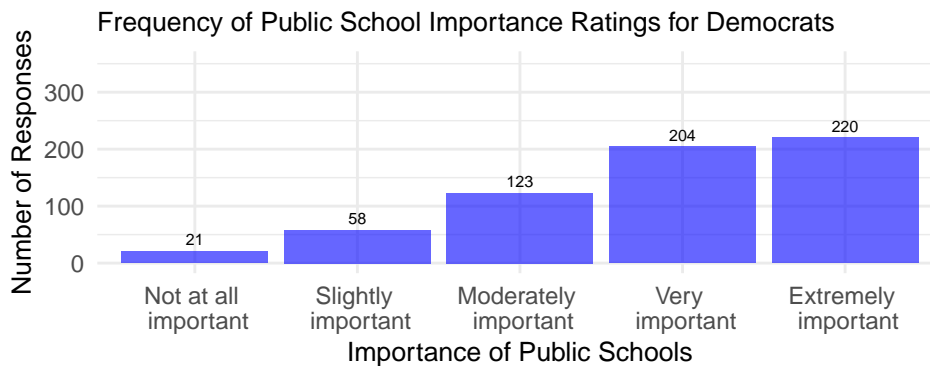
College professors



To first understand the data behind the distribution of the `group_colprofs` results, we plotted the distribution for both parties. This allowed us to see that while neither distribution is a direct t-normalized distribution they follow similar patterns that should allow us for evaluating other statistics. **The null hypothesis** is that the average temperature rating of Democratic voters towards college professors is equal to the average temperature rating of Republican voters. The alternative hypothesis is that average temperature ratings of Democratic and Republican voters towards college professors are not equal. In order to accurately use Welch's t-test, the data must fit the following assumptions:

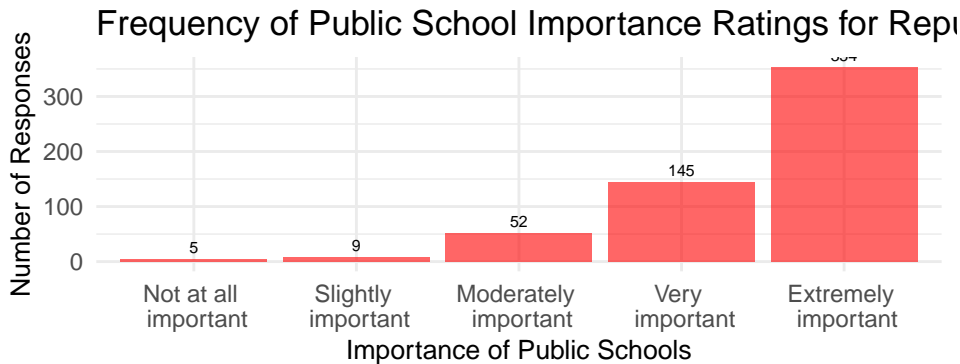
Independent and identically distributed (IID)? Participants were selected from the YouGov platform, where they receive an invitation to complete a survey and receive an award if the survey is completed. Although this poses issues for the independence of the data, since there could be relationships between participants who recommend the platform to each other, there are millions of users on the platform and therefore the related groups should have a minimal effect on the data. Additionally, YouGov performs sample matching to ensure that the ANES sample matches the distribution of differences amongst American citizens by pooling together public data from various surveys. This means the data is not quite as identically distributed as a random sample would be. However, ANES then weights the data to overestimate the standard error such that the identical distribution of the data should not present significant discrepancies in the interpretation of results. **Metric scale?** The data from college professor ratings is on an interval scale, where the order of the values is known (from 1-100, increasing) and we can quantify the difference between the steps on the scale, in this case the difference is 1 between each step. Therefore, our data is metric. **Relatively normal distribution given the sample size?** After filtering down to only Democrats and Republicans who voted, we have a total sample size of 676. This sample should be sufficiently large for the Central Limit Theorem (CLT) to kick in, meaning that the distribution will follow a mostly normal curve.

Importance of Content Taught



Warning in `plot_theme(plot)`: The ``plot.tt`` theme element is not defined in the

element hierarchy.



For the `imp_schteach` variable, we created two bar charts displaying the number of responses for each Likert category among Republicans and Democrats. The visualizations show that responses are more concentrated in the higher Likert categories for both Democrats and Republicans. **The null hypothesis** is that the probability that a draw from the Democratic Party ranks higher than a draw from the Republican Party is the same as the probability that a draw from the Republican Party ranks higher than a draw from the Democratic Party. The alternative hypothesis is that the probability that a draw from the Democratic Party ranks higher than a draw from the Republican Party is different than the probability that a draw from the Republican Party ranks higher than a draw from the Democratic Party. Let $p < 0.05$ be the rejection criterion for the null hypothesis. Since `imp_schteach` is ordinal and unpaired, the most suitable test for the data is the Wilcoxon rank-sum test (hypothesis of comparisons version). To perform this test, the data must meet two important assumptions:

Ordinal scale? If the data is not ordinal for a Wilcoxon rank-sum test (hypothesis of comparisons), the rank-based comparison of the null hypothesis loses meaning, and the test would not be able to determine if there is a difference between how Democrats and Republicans rank the importance of what's taught in schools. In this case, the data is ordinal because the response categories follow a meaningful order, but the differences between them are not necessarily equal. **Independent and identically distributed (IID)?** Violating independence can inflate observed differences, leading to artificially small p-values and an increased risk of Type I errors. Likewise, if identical distribution is violated due to weighting or differing group variances, the test may misrepresent the true effect. Given that the sample is drawn from a panel of 1M+ respondents, direct interdependence between respondents is unlikely to substantially inflate Type I/II errors. However, the opt-in nature of the panel may introduce dependencies such as shared characteristics between respondents. Additionally, since the weights adjust the influence of individual responses, the dataset may not be identically distributed, as some observations contribute more to the overall analysis than others. However, this corrects for under-representation and mitigates potential systematic bias.

Results and Interpretation

College Professors

```
Welch Two Sample t-test

data: group_colprofs by party_cat
t = 20.627, df = 999.49, p-value < 2.2e-16
alternative hypothesis: true difference in means between group Democrat and group Republican is not equal to 0
95 percent confidence interval:
 27.27697 33.01245
sample estimates:
 mean in group Democrat mean in group Republican
      71.36453           41.21982
```

After conducting Welch's t-test, we see that the difference in means is not equal to zero, with a P value of $2.2e-16$. This small p-value tells us that, under the assumption that the null hypothesis were true, there is a very small chance that we would have observed a difference as large as we have. Additionally, the confidence interval (27, 33) does not include zero, showing that we would not expect to find a difference of means equal to zero in our data. Practically speaking, the test results show that there is reason to believe there is a large difference between how Democrats and Republicans view

college professors, with Democrats rating them higher than Republicans overall. We can use Cohen's D to calculate the practical significance of this effect size.

When calculated, we get $d = 1.30$, and given that 0.8 is generally accepted as a large effect size we can know that our effect size is substantially large in comparison. This tells us that there is reason to believe that Democrats and Republicans have significantly different views on college professors, and leads us to consider what can be done to bridge the gap in understanding what these differences are.

```
d      |      95% CI
-----
1.30 | [1.17, 1.44]
```

Warning in rm(df): object 'df' not found

Importance of Content Taught

```
Wilcoxon rank sum test

data:  imp_schteach_num by party_cat
W = 119138, p-value < 2.2e-16
alternative hypothesis: true location shift is not equal to 0
```

From the results, we see that the p-value is extremely small (close to 0). Since $p < 0.05$, we reject the null hypothesis, suggesting that the distributions of perceived importance of what's being taught in public schools differ between political groups. While statistical significance tells us that a difference exists, we must assess practical significance to determine whether the difference is meaningful in context. We can use Cliff's Delta to assess this difference, which is well-suited for comparing ranks between groups.

```
Cliff's Delta

delta estimate: -0.3263168 (small)
95 percent confidence interval:
  lower      upper
-0.3815616 -0.2687492
```

Let D be a random variable representing the rank of a randomly selected Democrat. Let R be a random variable representing the rank of a random selected Republican. We can see that the delta, $P(D > R) - P(R > D)$, is roughly -0.326. The effect size is considered small, falling close to the bound between small and medium (0.147, 0.33). The value suggests that Republicans tend to rank the importance of public school education higher than Democrats, since the $P(R > D) > P(D > R)$.

Discussion

By examining Democrat and Republican perceptions of college professors and the importance of what's being taught in schools, we gain valuable insights into the broader American perspective on primary and secondary education and uncover partisan differences in ratings. Our analysis reveals a substantial partisan gap in how Democrats and Republicans perceive college professors, with Democrats generally assigning higher ratings. The large effect size in these ratings suggests a deep ideological divide in perceptions of higher education. Additionally, we found notable differences in how each group perceives the importance of what is being taught in schools, with Republicans generally ranking this importance higher than Democrats. These findings indicate that partisan differences are statistically significant in the areas we explored. However, further analysis of additional education-related variables is necessary to fully address the broader research question. Despite these differences, it is important to note that both parties agree on the moderate to high importance of public schooling, revealing a shared priority for the education system. These findings highlight the critical need to understand partisan perspectives in order to effectively inform education policy discussions and bridge ideological divides in the future.