Lab 1 by the Mean Machine

https://github.com/mids-w203/lab-1-mean_machines

Ender Ricart | Kelvin Yang | Alex Kim

March 4, 2025

Introduction

The US education system—from elementary school through graduate programs—depends heavily on federal money. Currently, with Republicans controlling Congress, education funding has been reduced. Recent Pew Research Reports on differing views on education between Republicans and Democrats (2018, 2022), found most Republicans viewed K-12 public education negatively and regard colleges as having a negative effect on the United States. Our research aims to contribute to a data-driven discussion of partisan educational politics and its real world impact on the public and private funding of our colleges and universities that have come to define our nation's thought leadership. In this study, we analyze data from the American National Election Studies (ANES) 2024 Pilot Study to understand: Do Republicans and Democrats have different views on higher education?

Data and Methodology

The data used in this study comes from the American National Election Studies (ANES) 2024 PIlot Study launched to the YouGov Internet survey platform and participant panel for completion. ANES recommends that the sample be described as follows: The survey was conducted using non-probability sampling. This method produces a sample that looks similar to a probability sample on the matched characteristics but may still differ in unknown ways on unmatched characteristics. Any analysis intended to make population inferences should be weighted. The study includes 1,500 weighted cases that passed quality control and were selected through sample matching to represent the population, plus 409 unweighted cases, included only for methodological analysis. Given our research focus on higher education, we identified two key variables within the survey to perform statistical analysis.
1. *group_colprofs* asks respondents to rate how they feel about college professors. We considered the demographic difference that more Democratic voters report some college education (n = 441) compared to Republican voters (n = 330), but determined this difference was not substantial enough to warrant additional data manipulation. For this variable, we conducted a Two-Sample T-Test, as the 0-100 thermometer scale provides appropriately granular interval data for metric-based testing.
2. *highered_approve* asks respondents to directly rank their approval of how colleges and universities are run. Our statistical analysis compares Democrat and Republican responses to this question using a Wilcoxon Rank-Sum test, which is appropriate for the ordinal Likert scale data.

Who is a "Voter"? Who is "Republican" or "Democrat"?

An important step in our study before conducting tests was preparing the data for exploration and analysis. Namely, we must classify someone as a Democrat or Republican, despite the sometimes ambiguous or contradictory question response about an individual respondent's partisan lean. Our approach first cut down the sample population to include only those who were registered to vote [*votereg* == "Yes"], leaving a sample population of 1,615. Within this sample we grouped people as "Republican" or "Democrat" if they responded to any of the following three question types in the affirmative "Republican" or "Democrat":

1. *pid__pid1d/pid__pid1r*: (Randomized to have Democrat or Republican listed first) Generally speaking, do you usually think of yourself as a Democrat(alt. Republican), a Republican (alt. Democrat), an independent, or what?
2. *pid__partylean*:(Randomized to have Democrat or Republican listed first based on pid_pid1) Do you think of yourself as closer to the [Democratic/Republican] Party or to the [Republican/Democratic] Party?
3. *pid3*: Generally speaking, do you think of yourself as a (Democrat, Republican, Independent, Other)?

We did not include "pid_7: Profile variable: 7 point party identification", "pid_dstr Strong Democrat?", "pid_rstr Strong Republican?" as these questions only added noise in the form of degree of party association ("lean", "strong", "not very strong"), without impacting the reported affiliation of Democrat or Republican. Our resulting group totals were 605 Democrats and 525 Republicans among registered voters.

Results: Rating College Professors (*group__colprofs*)

The survey asks respondents to rate college professors on a 'feeling thermometer' scale from 0 to 100, with 50 being a neutral score. We sought to use these responses to understand whether Democrats and Republicans differ significantly in their views of college professors. Of 605 identified Democrats in the survey, 66 declined (selected 'Next') to answer or were not asked this question; of 525 identified Republicans, 40 declined to answer (selected "Next') or were not asked this question. In our analysis, we excluded these responses.

To analyze the data, we selected a Two-Sample t-Test to test our **null hypothesis:** *Democrats and Republicans share the same views on college professors.* More specifically for the variables under consideration, our Null Hypothesis is as follows: *The expectation of democratic ratings of college professors is equal to the expectation of republican ratings of college professors (the difference of the means is 0).* A statistically valid Two-Sample t-Test requires three assumptions:

1. Metric scale data: Although the survey question asks respondents to rate college professors, the response is framed on a 'thermometer' scale providing sufficient granularity in reponse values. It offers a familiar metric framework through association with a common measurement tool that respondents can interpret consistent and equal intervals between 0 to 100.

2. I.I.D. data: The study sampled from the population of US citizens who are at least 18 years old. The study was conducted by several prominent University affiliates, including University of Michigan and Stanford. Respondents were selected for their representation of population diversity. We assume that their sampling method is sufficiently random to support the assumption of independence. Although the survey is opt-in, we also determined that the samples are sufficiently drawn from the same distribution.

3. No major deviations from normality: The histograms (see Appendix Figure 1.1 and Figure 1.2) reveal that the data is somewhat skewed, particularly for Democrat responses. We know from our data exploration that more Democrats than Republicans have college-level or more education which may contribute to the skew. However, given the substantial sample

size (n >= 330 for both groups), we can rely on the Central Limit Theorem to satisfy the assumption of normality for this test.

```
t.test(d_colprofs, r_colprofs)
cohensD(d_colprofs, r_colprofs)
```

With a p-value less than 2.2e-16 from our Two-Sample t-Test, we **reject the null hypothesis**. Given such a small p-value, we are confident in the interpretation of our results. To better understand the practical significance, we computed Cohen's D, which measures the difference between two sample means in standard deviation units, to be 1.12. This indicates a large effect size for the difference, highlighting the gap in attitudes between Democrats (tending to view professors positively) and Republicans (tending to view professors negatively). In the next section, we aim to understand whether this gap extends to opinions of how higher education institutions are run generally.

Results: Ranking How Universities are Run (*highered_approval*)

The survey includes questions about respondents' approval of how colleges/universities are run. Data is reported on a 5-point Likert scale that ranges from 1 ("Approve strongly") to 5 ("Disapprove strongly"), and the same question is asked to each respondent. There were 168 out of 1,909 total responses marked as "inapplicable, legitimate skip," which indicates that the respondent did not receive it because the question was irrelevant to the respondent based on their previous answer. These responses were excluded from the analysis to ensure that the study included only respondents who directly answered this question.

Using this response, we aimed to determine whether Democrat's and Republican's approval of colleges and universities differ significantly. We selected a Wilcoxon Rank-Sum Test to test our **null hypothesis**: *Democrats and Republicans share the same views on higher education.* More specifically for the variable under consideration here, our Null Hypothesis can be stated as follows: *The probability that a randomly selected Democrat ranks higher than a randomly selected Republican is the same as the probability that a randomly selected Republican ranks higher than a randomly selected Democrat.* For a Wilcoxon Rank-Sum Test to produce reliable inference, the data must, at minimum, be ordinal, and the two groups must be i.i.d. First, since the question asks each respondent to answer using the provided Likert scale, the data is ordinal and meets the first assumption for this test. Second, each respondent answered the same survey questions independently and was classified into Democrat and Republican based on their self-report party affiliation. Along with the details from our previous assumption write up, this too meets the i.i.d. assumption.

```
wilcox.test(bias_d$highered_approve, bias_r$highered_approve)
```

The Wilcoxon Rank-Sum Test found a highly significant difference (p < 2.2e-16) in the approval of colleges and universities between Democrats and Republicans. Therefore, we **reject the null hypothesis**.

4

We used the Common Language Effect Size (CLE) to assess the magnitude of the difference. The CLE result was 0.208, meaning there is a 20.8% probability that a randomly selected Democrat would rate higher (i.e., more disapproving) than a randomly selected Republican. Conversely, there is a 79.2% probability that a randomly selected Republican would rate higher (i.e., more disapproving) than a Democrat. This indicates that Republicans were significantly more likely to disapprove of higher education institutions than Democrats, reflecting a strong partisan divide. The bar charts (see Appendix Figure 2.1 and 2.2) also support this result. The distribution shapes differ significantly, with Democrats being more neutral or approving and Republicans being largely disapproving. These findings highlight deep partisan divides in perceptions of higher education institutions.

Discussion

The significant differences in how Democrats and Republicans rate college professors and view higher education institutions confirm a substantial partisan divide that extends beyond mere political disagreements into fundamental attitudes toward academia. This polarization has potential implications for federal education funding policies, institutional support, and the broader societal value placed on higher education in America. These findings suggest that educational institutions may need to develop strategies to bridge this partisan gap if they hope to maintain broad-based public and political support across the ideological spectrum.

Appendix

Figure 1.1

Figure 1.2

Figure 2.1

Figure 2.2