Data Analysis for Cornell Votes

Adriana Lorena Jimenez Bonilla

2024-12-01

Introduction

This project focuses on analyzing and communicating insights from Cornell's NSLVE (National Study of Learning, Voting, and Engagement) reports. These reports provide data on student voter participation, serving as a foundation to strategize and enhance voting efforts on campus. The ultimate goal is to ensure Cornell's voting rate remains at or above the average voting rates of similar institutions in NSLVE. In this case, similar institutions consist of private research institutions.

Data Description

The data analyzed in this project comes from Cornell University's NSLVE (National Study of Learning, Voting, and Engagement) reports for the years 2012-2020. These reports provide detailed insights into student voter participation, including registration rates, voting rates, and voting methods. The data is disaggregated by demographic and academic characteristics such as age, gender, field of study, and enrollment status. Adjustments have been made to exclude ineligible voters, such as non-resident aliens, based on institutional records. The reports compare Cornell's performance to averages across similar institutions, which is where our primary focus lies.

The objective of this project was to process and harmonize voting-related datasets across multiple election years to enable the comparison of demographic statistics between 2012 and 2022. Two datasets containing 2016 and 2018 data contained conflicting but similar values. To resolve these inconsistencies, we averaged the numeric columns by computing the mean of corresponding values.
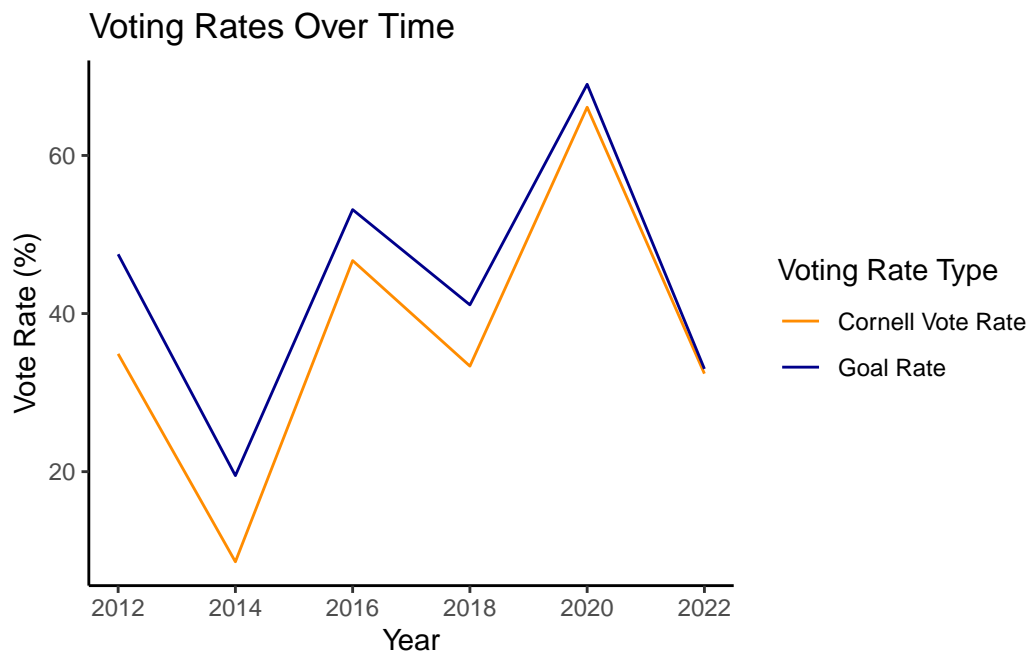
To prepare for combining all datasets into a single dataframe, we added a new column named `year` to each dataset to indicate the corresponding election year. This ensured the data could be differentiated after merging. The unified dataset is now ready for further analysis, such as comparing voting rates, registration rates, and demographic trends over time.
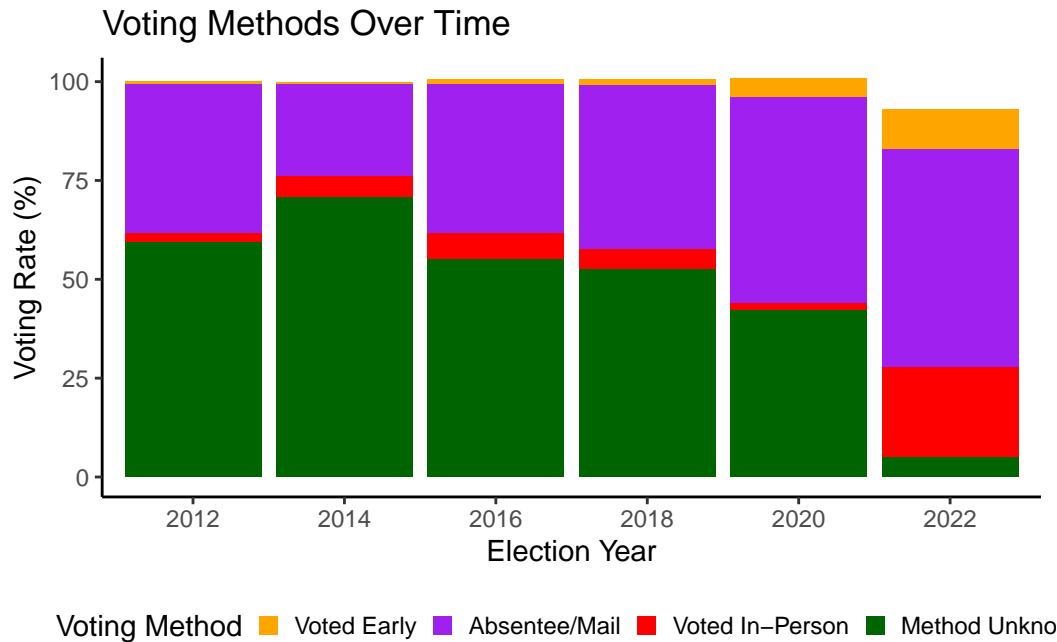
A description of each variable (column) is as follows:

- corn_reg_rate: the percentage of voting-eligible students who registered to vote.
- corn_vote_rate: the percentage of eligible students who voted on Election Day.
- goal_rate: the voting rate of all research institutions, our goal voting rate.
- method_nonperson: the percentage of people who voted absentee/by mail.
- method_early: the percentage of people who voted early.
- method_person: the percentage of people who voted in person.
- method_unknown: the percentage of people whose voting method is unknown.
- age1: the voting rate of people between ages 18-21.
- age2: the voting rate of people between ages 22-24.
- age3: the voting rate of people between ages 25-29.
- age4: the voting rate of people between ages 30-39.
- age5: the voting rate of people between ages 40-49.
- age6: the voting rate of people aged 50+.
- ugrad: the voting rate of undergraduate students.
- grad: the voting rate of graduate students.
- fresh: the voting rate for first-year undergraduates (freshmen).
- soph: the voting rate for second-year undergraduates (sophomores).
- up: the voting rate for upperclass undergraduates (juniors and seniors).
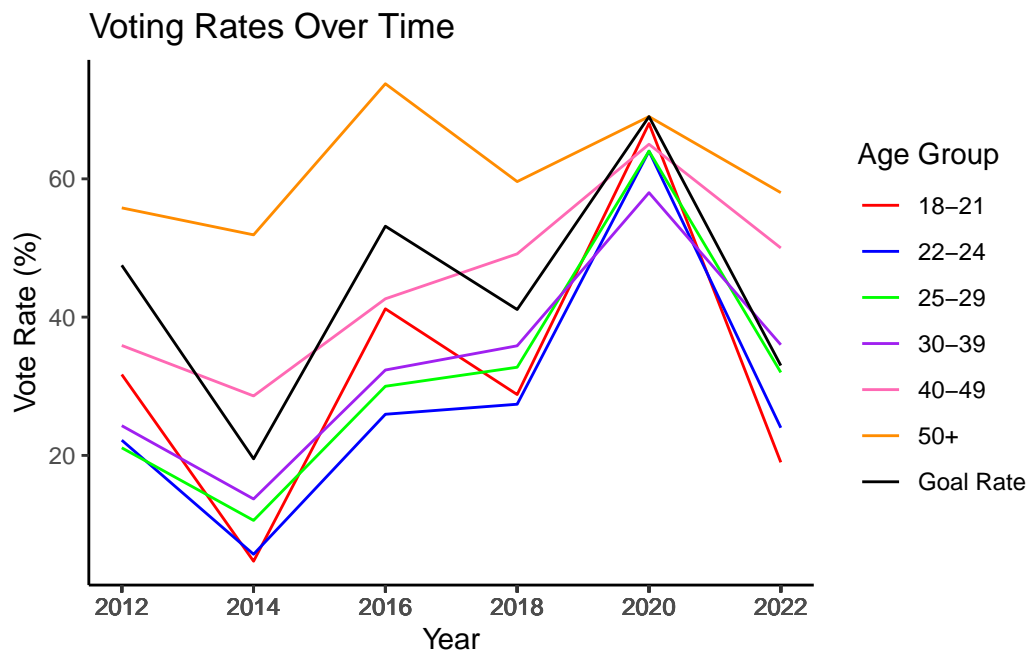
Data Analysis

Research Question: How can Cornell University sustain or improve its voter turnout rates to remain at or above the average voting rates of private research institutions?

2

Voting Methods Over Time

Voting Rate (%) — 0, 25, 50, 75, 100

Election Year — 2012, 2014, 2016, 2018, 2020, 2022

Voting Method: Voted Early ■ Absentee/Mail ■ Voted In–Person ■ Method Unkno ■

While Cornell Votes makes an active effort not to persuade prospective voters toward any one voting method, it may be beneficial to consider trends among students so we can best support whatever methods they choose.

Voting Rates Over Time

Vote Rate (%) — 20, 40, 60

Year — 2012, 2014, 2016, 2018, 2020, 2022

Age Group:
- 18–21
- 22–24
- 25–29
- 30–39
- 40–49
- 50+
- Goal Rate

The only group consistently at or above our goal rate is voters aged 50+. Second best, the group of voters ages 40-49, is interestingly above our goal during midterm elections while below our goal during presidential elections. If we focus on the remaining population, voters age 18-21 are consistently the most active during presidential elections, but have the steepest decline during midterm elections. Groups 22-24, 25-29, and 30-39 saw increases between 2014-2020, but dropped again in 2022. How can we make people care about midterm elections, and make that care everlasting?
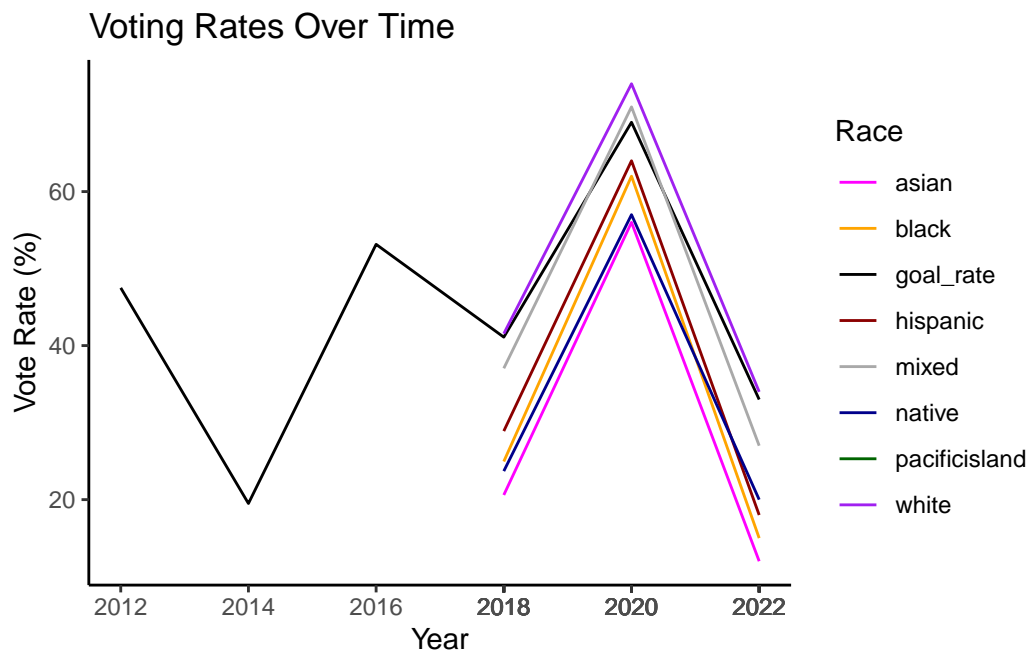
Voting Rates Over Time
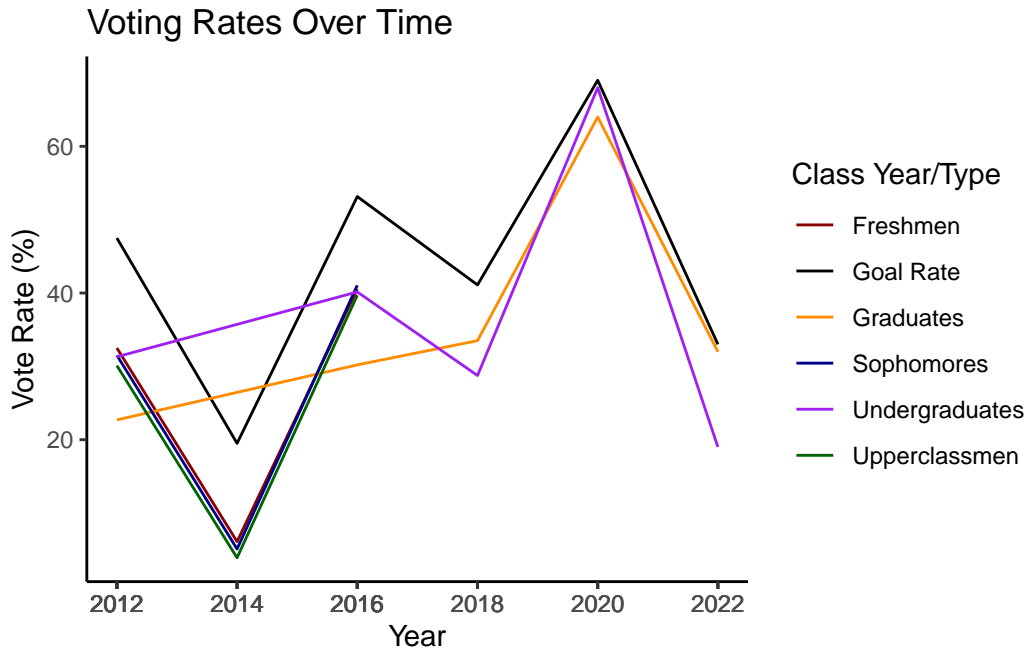
While some data was unavailable, making the graph incomplete, we can still note a few things. We see that, from 2012-2016, freshmen, sophomores, and upperclassmen moving in nearly identical trends. Then, from 2016-2022, we see that changes in undergraduate student voting rates are more drastic from year-to-year than that of graduate students. Notice that from 2016-2018 undergraduate voting rates seem to diverge from our goal rate, while that of graduate students converges towards it. Why might that be?

The two previous graphs reflect two key things: 1) young voters are more impressionable (i.e. more susceptible to changes in Cornell Votes' curriculum) and 2) young voters are less involved in midterm elections. Since voting is habit-forming, it may be important to consider why the spike in 2020 was followed by a consistent plummet across all ages in 2022. This presents the question, how can we retain high voting rates from year to year (i.e. how can we make people care about midterm elections)?

4

Voting Rates Over Time
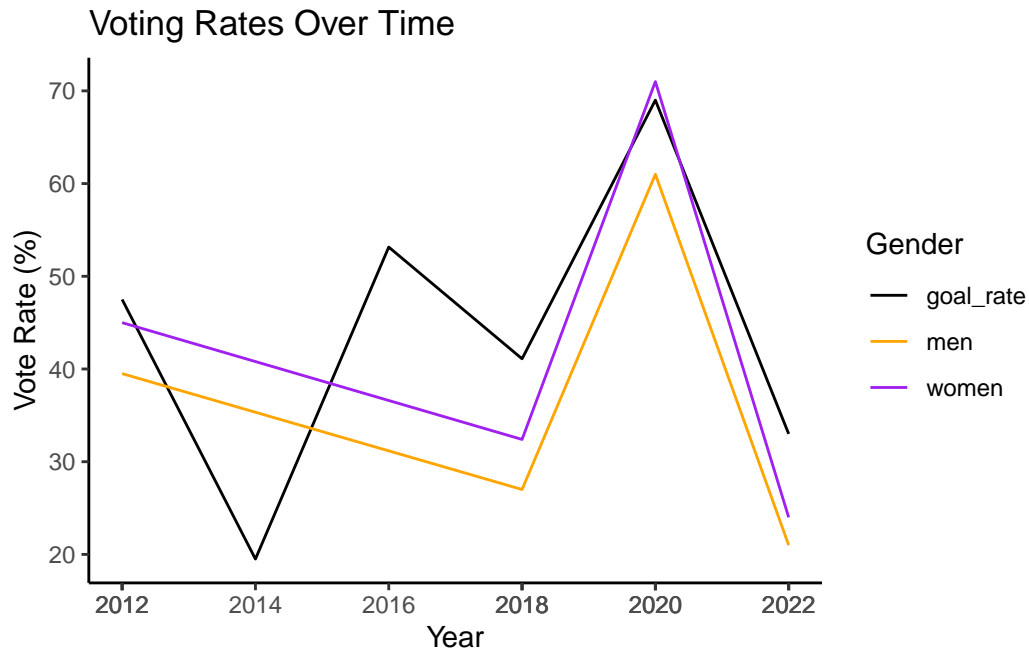
While the trends are indentical, women consistently present a higher voting rate than men. Women actually surpassed our goal in 2020, before falling short of it in 2022. Men consistently present numbers below our goal.

Voting Rates Over Time

While the trends are nearly identical, white students appear to vote above our target rate. How can we encourage minority students to vote?

Average Field of Study Voting Trends Over Time
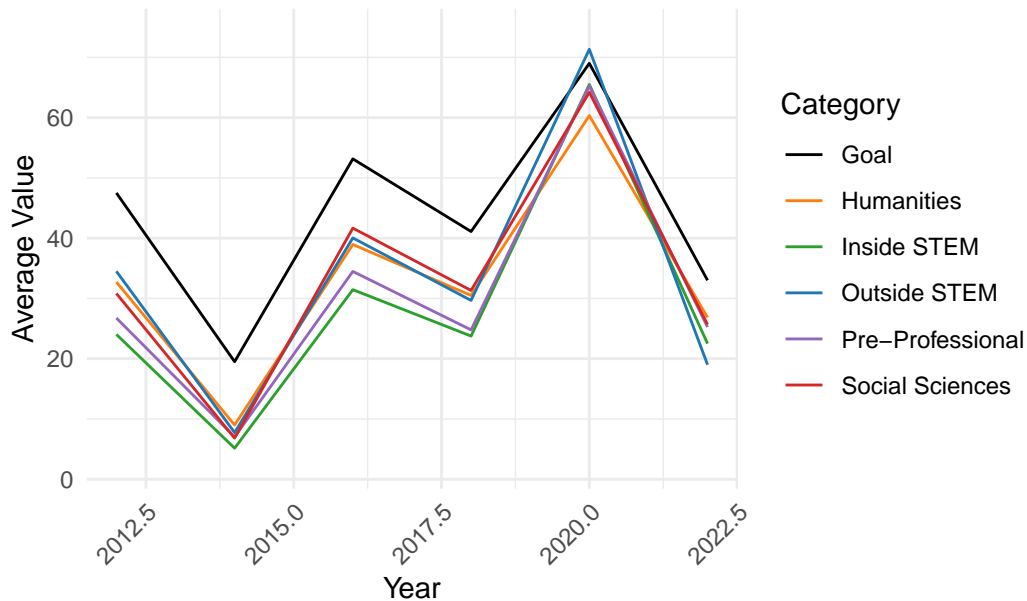
The primary dataset used in this project represents the average values of students' participation in various fields of study at Cornell University over multiple years. To simplify the analysis and highlight overarching trends, I categorized the individual fields of study into broader, more manageable groups based on their similarities and common themes. These categories reflect key academic disciplines and their intersections, helping us to observe trends in educational patterns and make comparisons across different areas of study. Here's how I divided the fields of study into categories:

- Humanities: History, Philosophy and Religious Studies, Liberal Arts and Sciences and Humanities, Communication and Journalism.
- Inside STEM: Biological and Biomedical Sciences, Computer and Information Sciences, Engineering and Engineering Technologies, Mathematics and Statistics.
- Outside STEM: Agriculture, Natural Resources and Conservation, Physical Sciences.
- Pre-Professional: Architecture, Business Management and Marketing, Legal Professions and Studies, Health Professions.
- Social Sciences: Area, Ethnic, Cultural, and Gender Studies, Family and Consumer Human Sciences, Psychology, Public Administration and Social Services, Social Sciences.

Looking at the generalized disciplines above, we see that general trends are similar, but some groups are consistently less than or greater than others. The only group who performed

above our goal in any given year was `Outside STEM` in 2020. Let's look at how others have diverged/converged. Do we notice anything?

Key Takeaways

Ideas for the Future