Creating the Dataset

• Our institution submits enrollment records to the Clearinghouse throughout the year. Which enrollment record do you use?
  We use the record sent by your institution to the Clearinghouse on the closest date before the election.

• What data fields from student enrollment records are used for matching?
  Clearinghouse uses name, date of birth, and address on file to match enrollment records to public voting files. Once the data are matched, the Clearinghouse removes all names, dates of birth, and addresses, as well as any other information that would allow IDHE researchers to identify an individual student before sending IDHE the file. IDHE receives a unique student identifying number, an age on the date of the election, and a zip code of the address on file.

• What data fields from student enrollment records are sent to IDHE?
  IDHE receives the matched file from Clearinghouse after they have made the records confidential. In addition to removing personally identifiable information (PII), Clearinghouse also redacts non-PII when record counts are small. We receive the following student-level data from enrollment records:
  - Age at Election – Clearinghouse sends IDHE a computed variable indicating the student’s age on the day of the election.
  - Class level
  - Field of Study
  - Zip code for the address on file
  - Race/ethnicity (if provided to Clearinghouse by the campus)
  - Sex (if provided to Clearinghouse by the campus)
  - Nonresident alien status (if provided to the Clearinghouse by the campus)
  - Enrollment Status (full/part time)
  - Degree Seeking Status (is/is not degree seeking)

• Where can I learn more about the matching methodology?
See our fall 2020 methodology update, which describes our matching process and post-match processing.

Campus Report Contents

- You provide three topline rates in the campus report. Please explain the differences.
  The Voting Rate is the percent of voting-eligible students at the institution who voted an election. The voting-eligible enrollment population is the total number of enrollment records that we receive from the Clearinghouse, less the number of students under 18 at the time of the election and estimated number of non-citizen students.
  The Registration Rate is the percent of voting-eligible students who registered to vote prior to an election.
  The Voting Rate of Registered Students or “Yield Rate” is the percent of registered students who then voted in an election.

- What does “*” (asterisk) mean in a graph?
  An asterisk indicates that the number of students represented in this data point is fewer than ten. We do not report information with small cell sizes to protect student privacy.

- What does “-“ (dash) mean in a graph?
  A dash indicates that there are zero results in the field.

- Some graphs contain information for “unknown” students. What does that mean?
  Your institution is probably reporting to the Clearinghouse that number of students without additional data (e.g., race, gender, or class level). You can improve the quality and usefulness of your report by submitting missing data to the Clearinghouse. Vote method data comes to us from the voter file, so its availability depends on state and district reporting. If a state or district does not maintain vote method, then we cannot include it in our calculations. Currently, vote method information is unavailable for students who voted in Alabama, Delaware, Idaho, Illinois, Kansas, Kentucky, Louisiana, Massachusetts, Missouri, Mississippi, Nebraska, New Mexico, New York, Ohio, South Carolina, South Dakota, and Vermont.

- How did you determine fields of study?
  The Clearinghouse uses IPEDS CIP information to cluster more than 8,100 discrete fields into 51 Study Families. We are examining ways to process these data further to make them useful to our campuses.

Tabulations and Calculations

- In your national report, Democracy Counts 2018, you provide two voting rates, the Average Institutional Voting Rate (AIVR) and a National Student Voting Rate (NSVR). What is the difference, and which do you use in the campus report for institutional comparisons in 2018?
  The AIVR is the average of institutional voting rate – it answers the question “at an average campus, what is the voting rate?” This is the rate that would be comparable to an individual institution’s voting rate. The NSVR is the voting rate of all students at NSLVE-participating institutions.

- Where did you get the “total student enrollment” number in the report?
This is the number of records your institution submits to the Clearinghouse.

- **How do you know how many of our students are under the age of 18?**
  For each (de-identified) student record, we receive an age on the date of the election. We remove anyone under the age of 18 from the voting-eligible student count.

- **How do you know how many students have opted to exercise a FERPA block?**
  Your institution sends that information to the Clearinghouse, which then sends it on to us.

## Data Accuracy

- **Our IPEDS enrollment total is different from the Clearinghouse total. Why?**
  Usually, the registrar submits data to the Clearinghouse and an office of institutional research or assessment submits data to IPEDS. Clearinghouse data is individual level, whereas IPEDS data is reported in the aggregate. They may be counting students differently (e.g., part-time, degree seeking, international). You should check with the offices submitting the data to understand (and reconcile) the differences.

- **How accurate are these calculations?**
  Your estimated voter registration and voting rates are based on enrollment records that are matched to publicly available voting records. This process removes a large source of error inherent to most voting estimates, which usually rely on self-reported responses to surveys or estimates without a firm number of eligible voters. That said, NSLVE is both a new project and new process. We want to collaborate with representatives at participating institutions to flag and work with us to correct problems. Since the first set of reports were issued in 2014, we have identified several data issues that we then correct through testing and improving the matching process and campaigns to campuses to improve student-level data. Your estimates improve with each election, but we do rely on you to flag things that don’t look right.

We have identified two main categories of issues: (1) false positive or negative results from the matching process and (2) incomplete enrollment records submitted by campuses to the Clearinghouse.

### Issues Related to Matching Enrollment Records to Voting Records

Errors in the matching process can be more difficult to correct due to our system of building the database. IDHE never receives personally identifiable information (PII) or student names, and as a result, we cannot cross validate the results of the matching process with the relevant public voting records. Enrollment data are matched to voting records using a system developed by the external vendor that compiles the voting records. This matching process is a sophisticated algorithm, but it’s not perfect and might lead to false positive or false negative matches. Enrollment records that do not match voting records are usually nonvoters (i.e. they do not match because there is no record to match), but the matching algorithm may fail because the data fields are inaccurate – for instance, a student’s name in an enrollment record may not match the way it is present in a voting record—or because the campus location data (city, state, and zip code) does not align with the address used by students when they register to vote (e.g., “PO Box 123, Tufts”). We are currently testing a new matching process and will keep you posted on issues across the varied geographies and institutional types among NSLVE campuses.
Problems Related to Incomplete Enrollment Records

Incomplete or inaccurate enrollment data submitted by campuses to the Clearinghouse has long been a problem. While the data is improving, we need more support from campuses. For this reason, we added institutional researchers as report recipients in the 2018 dissemination process in the hope that reporting errors will be caught and corrected. We use the campus-submitted Clearinghouse data to determine the student enrollment and the number of students ineligible to vote by age (under 18 at the time of the election). We use IPEDS to estimate how many students are non-resident aliens (often international students) who are ineligible to vote because of citizenship status. Ideally, IPEDS and Clearinghouse totals match.

At this time, we can identify a few common causes of record incompleteness:

1. **Some campuses do not report all students to the Clearinghouse.** One source of confusion might be whether to use Social Security numbers as the reporting unit. At one time, Clearinghouse submissions required a Social Security number; students without SSNs on file could not be submitted. That rule changed in 2009, but a few campuses continue to submit only students with SSNs. It is possible that some students who are eligible to vote nonetheless do not provide an SSN to the institution, and this is more likely to be true of students not receiving financial aid. This will bias the voting rate estimate because the omitted student records are not random. Similarly, some campuses only submit to the Clearinghouse students who are US citizens. Noncitizen students are counted in the numbers for IPEDS, and our calculations assume they are present in the campus enrollment records. Campuses will improve the accuracy of their NSLVE reports if they include all students in the Clearinghouse enrollment file and flag nonresident aliens under the category for race/ethnicity.

2. **Some campuses report different enrollment numbers to IPEDS and Clearinghouse.** For example, under the same OPEID scheme some campuses report the full headcount of students as affiliated with one branch of an institution to IPEDS and submit an incomplete headcount of students to Clearinghouse, or vice versa.

3. **Some campuses define “students” differently.** A few campuses treat post-docs as students in their submission to Clearinghouse when, in fact, they are employees and should not be included in either submission. It is also important that students be identified as degree seeking or non-degree seeking in both the submission to Clearinghouse and IPEDS. We include only degree granting institutions in NSLVE, and track voting for both degree-seeking and non-degree seeking students, either full- or part-time. Complete data in these categories is important.

4. **Some campuses have unusually high numbers of records blocked by a FERPA request.** Some students block their directory information from any use, including for research, pursuant to the Family Educational Rights and Privacy Act (FERPA). Since we use directory information to perform the match to voting records, these students are removed from the enrollment records before the process of matching enrollment and voting records. About 4% of all students nationally exercise this right. We provide in each campus report the number of students removed due to a FERPA block. If that
number is large, please consider how students are presented with the option of blocking the use of their records (campuses requiring an affirmative decision by students have lower FERPA blocks). If those students voted at lower or higher rates, that could slightly affect the voting rate. They are also not removed from the aggregate count submitted to IPEDS, which could affect the enrollment numbers or estimated NRAs as well.

5. **Some campuses do not submit race/ethnicity data, which means we cannot easily remove non-resident aliens (mostly international students) from the database, nor can we provide campuses with data to show equity gaps between their students.** We have a workaround when institutions do not flag their international students. Institutions report to IPEDS their total enrollment and the number of non-resident aliens they serve. By calculating a percentage of non-resident aliens on each campus based on IPEDS number and applying that to the enrollment number you provide the Clearinghouse, we can estimate how many of your students are non-resident aliens. That, however, is a less-than-perfect work around and we would prefer to work with accurate numbers non-resident aliens. Note that we cannot remove these students from the disaggregated data starting on page 5 of your report. We have no workaround – and cannot report gaps in your students’ turnout rates - if campuses provide do not provide race/ethnicity and sex data to the Clearinghouse.

*Other Limitations*

We cannot remove from the database students who are ineligible to vote because they are resident aliens (i.e. green card holders and other noncitizens authorized to reside in the U.S. indefinitely) or have been disenfranchised due to a criminal conviction. Unfortunately, we have no way of adjusting for these students. There is no national database that keeps track of them, and the Clearinghouse does not collect that information. We do not expect campuses to compile and submit lists of undocumented or disenfranchised students, but we provide a [Recalculation Tool](#) on our website that you can use to recalculate your rates if you have these exact numbers.

- **How can we improve the accuracy of our reports?**
  Our accuracy is largely dependent on the quality of the data institutions submit to the Clearinghouse. Here are some of the more common problems with the data, and how to address them:

  1. **Provide race/ethnicity data and use the “nonresident alien” race designation:** Your reports will be more accurate if you provide the Clearinghouse with accurate information about a student’s nonresident alien status, which is a subset of the race and ethnicity category.
  2. **Submit all student records to the Clearinghouse:** In some cases, differences in enrollment totals are a result of inconsistencies in reporting part-time students. This is an easy problem to solve by indicating part- and full-time status for each student. Another valuable data element is degree- and non-degree seeking status.
  3. **Authorize IDHE as an exception to FERPA record blocks:** Nationally, 4-5% of students block the use of their names for any reason. If your number is higher than 5%, you might be able to change that with different instructions to students and clarification about how their names might be used (e.g., for research). The other option is to send to us an
authorization form that the Clearinghouse created in 2019 for institutions with high-FERPA-blocks. You can contact us at nslve@tufts.edu to discuss this process. With this new form, we can use FERPA blocked records under the “special studies” exception to FERPA.

You can address many of these issues by working with the office on campus, usually the registrar, that submits your data to the National Student Clearinghouse. The submission process is automated, so it may be a simple matter of asking the responsible person to change filters or add data fields. We suggest that you take the report over to the individual responsible for submitting information, show that person the number of fields marked with “unknown” in your report, and explain why this would be valuable. If you are missing entire pages, such as for race/ethnicity and gender, you might want to bring along a Sample Report that shows what you could receive. You may need to involve senior institutional leaders, in which case you should be prepared to explain how you and people at other campuses use the data.