Arizona School Board Elections & Accountability Data

Appendix: Data Description and Preliminary Descriptive Statistics

February 2026

1 Introduction

This appendix describes the construction, contents, and preliminary descriptive statistics for a dataset linking Arizona school board election results (2014–2024) with district-level accountability metrics from the Arizona Department of Education (ADE). The dataset covers six general election cycles and 138 of Arizona's 223 regular public school districts.

The two analysis-ready files are:

- **Master file** (`az_school_board_master.csv`): 2,984 candidate-level records with 36 variables, including vote totals, party affiliation, and district accountability data.
- **Summary file** (`az_district_year_summary.csv`): 394 district-year observations with 33 variables, including competitiveness measures, voter turnout, and aggregated accountability metrics.

2 Data Construction

2.1 Election Results

The primary election data were collected from Arizona county election offices for general elections in 2014, 2016, 2018, 2020, 2022, and 2024. These records were manually compiled from official canvass reports into an Apple Numbers workbook, then exported to CSV.

In a supplemental data collection pass, 461 additional candidate records (covering 95 district-year gaps) were ingested from the OpenElections Arizona repository, which provides county-level precinct results in machine-readable CSV format. This ingestion targeted districts already in the dataset that were missing coverage in specific election years. The largest gains were in 2024 (+49 district-years from Maricopa, Pima, and other county files) and 2014 (+36 district-years). A Python pipeline (`ingest_openelections.py`) automated the download, format detection, race identification, and district-name-to-CTDS matching.

2.2 District Identifiers

Each district is identified by its ADE County-Type-District-School (CTDS) Entity ID and cross-referenced with the NCES Common Core of Data (CCD) 2024–2025 Local Education Agency (LEA) file to obtain national NCES LEA IDs. The CTDS ID serves as the primary join key between election data and ADE accountability files.

2.3 Accountability Metrics

Five categories of ADE accountability data were merged onto the election records by CTDS ID and year:

- **A–F letter grades**: ADE A–F Accountability files (FY2013–14 through FY2024–25). Grades were not published in FY2015–16 or FY2020–21 due to a COVID-era moratorium.
- **Total enrollment**: ADE October 1 Enrollment counts (FY2014–2026).
- **Dropout rates**: ADE Dropout Rate reports (FY2013–2025), filtered to Subgroup = "All."
- **4-year graduation rates**: ADE Cohort Graduation Rates (Cohort 2013–2025), filtered to Subgroup = "All." Only available for districts operating high schools ($\sim$50% of districts).
- **Superintendent data**: Hand-collected from district websites, press releases, and LinkedIn (2025–2026 snapshot, in progress).

2.4 Summary File Construction

The summary file aggregates candidate-level records to one row per district per election year, computing derived analytical variables including number of candidates, number of seats, candidates-per-seat ratio, whether the race was contested, total votes cast, winner margin, and winner names. The summary excludes bond and budget override questions, county superintendent races, and rows without a CTDS ID.

3 Descriptive Statistics

3.1 Coverage by Election Cycle

The dataset spans six biennial election cycles. Table 1 reports the number of districts and candidate records in each cycle.

Table 1: Election data coverage by cycle

Year	District-Year Races	Candidate Records	Avg. Candidates/Race
2014	69	365	5.3
2016	86	674	7.8
2018	76	407	5.4
2020	51	378	7.4
2022	35	227	6.5
2024	77	933	12.1
Total	**394**	**2,984**	**7.6**

Figure 1 visualizes the year-to-year variation. The 2024 cycle has the most candidate records due to comprehensive OpenElections coverage of Maricopa County precinct data.
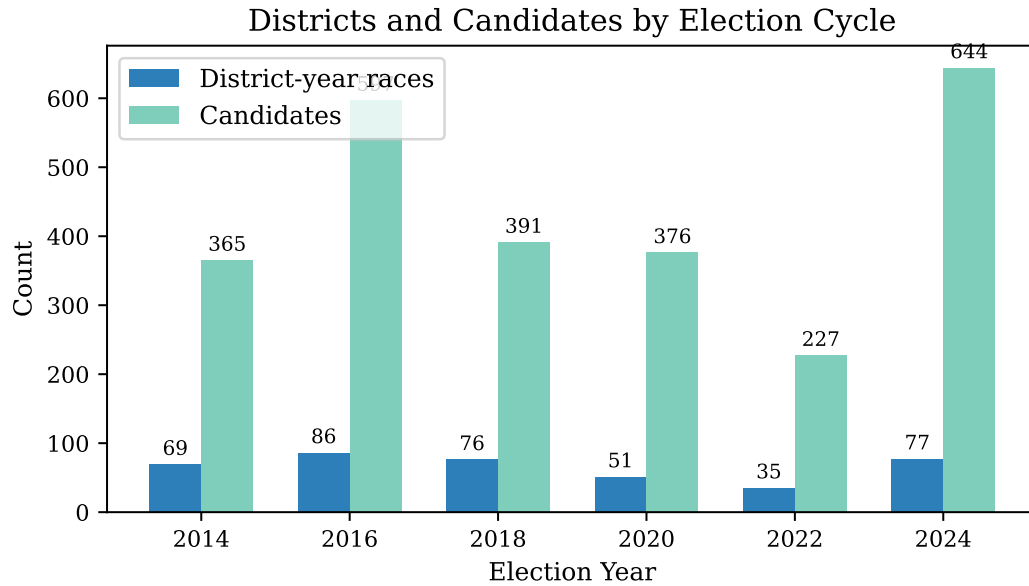
Figure 1: Number of district-year races and candidates by election cycle.

3.2 Geographic Distribution

The 138 districts span all 15 Arizona counties. Maricopa County accounts for the largest share of the data (1,410 candidate records across 54 districts), reflecting its position as the state's most populous county. Figure 2 shows the distribution of candidate records across counties.

3.3 District Size

District enrollment ranges from fewer than 100 students to over 130,000. Table 2 shows the size distribution of districts in the dataset. The five largest districts by enrollment are Mesa Unified (130,096), Tucson Unified (96,620), Gilbert Unified (75,865), Peoria Unified (74,323), and Deer Valley Unified (69,630).

Table 2: District size distribution (peak enrollment across years)

Enrollment Range	Districts
Under 500	23
500–2,000	22
2,000–10,000	48
10,000–50,000	36
Over 50,000	8

Figure 3 shows the full enrollment distribution, which is strongly right-skewed (median = 5,731; mean = 12,762).

3.4 Competitiveness

Of the 394 district-year races in the summary file, 286 (73%) were contested—meaning more candidates ran than seats were available. Competitiveness varies by year (Figure 4): 2016 (95%), 2018
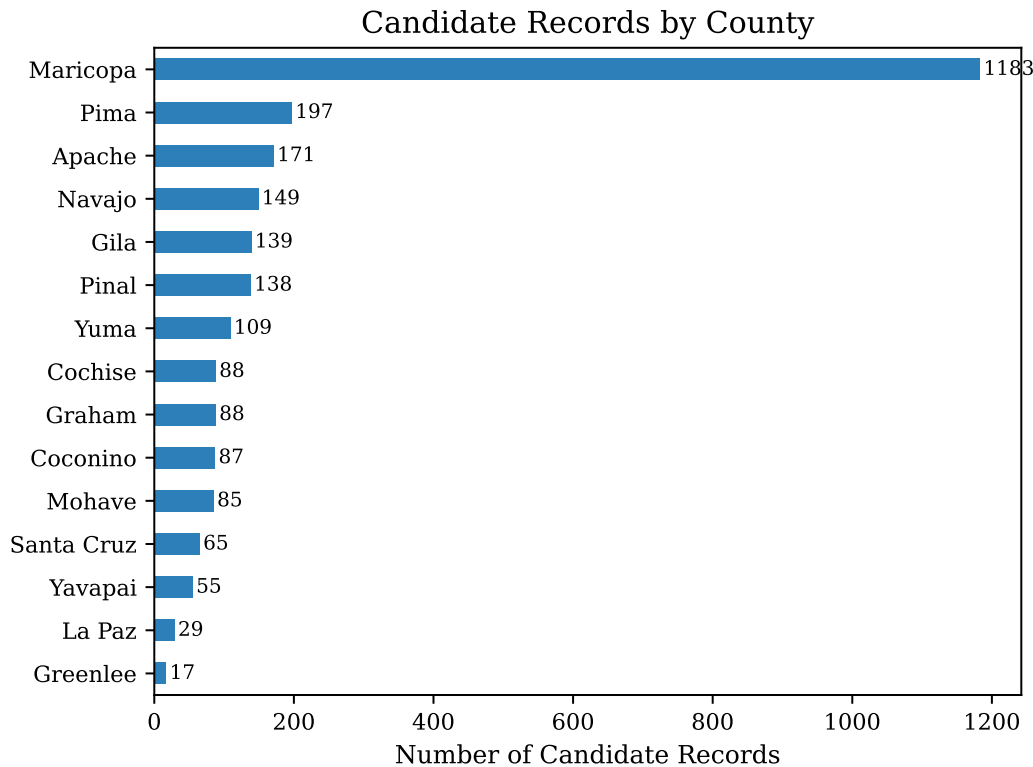
Candidate Records by County

Figure 2: Candidate records by county.

(92%), and 2022 (97%) had the highest contested rates, while 2014 (45%) and 2024 (36%) had lower rates. The lower rates in 2014 and 2024 are partly an artifact of the OpenElections-sourced records, for which seat counts were often unavailable, making it impossible to determine whether a race was contested.

The average number of candidates per seat is 3.8 across all races with seat data, ranging from 2.7 in 2014 to 4.9 in 2020. Figure 5 shows the distribution.

3.5   Winner Margins

Among contested races with margin data, the median winner margin is 3.0% and the mean is 8.3%, indicating that most races are relatively close. Figure 6 shows that the distribution is heavily right-skewed, with a long tail of blowout races.

3.6   Voter Turnout and District Size

Figure 7 plots total votes cast against district enrollment. As expected, larger districts attract more total votes, though the relationship is not strictly proportional—some large districts have relatively lower turnout, and some small districts have surprisingly high participation.

3.7   Party Affiliation

Party data is available for 1,200 of 2,984 candidate records (40%). Most Arizona school board races are officially nonpartisan. Among rows with party data:
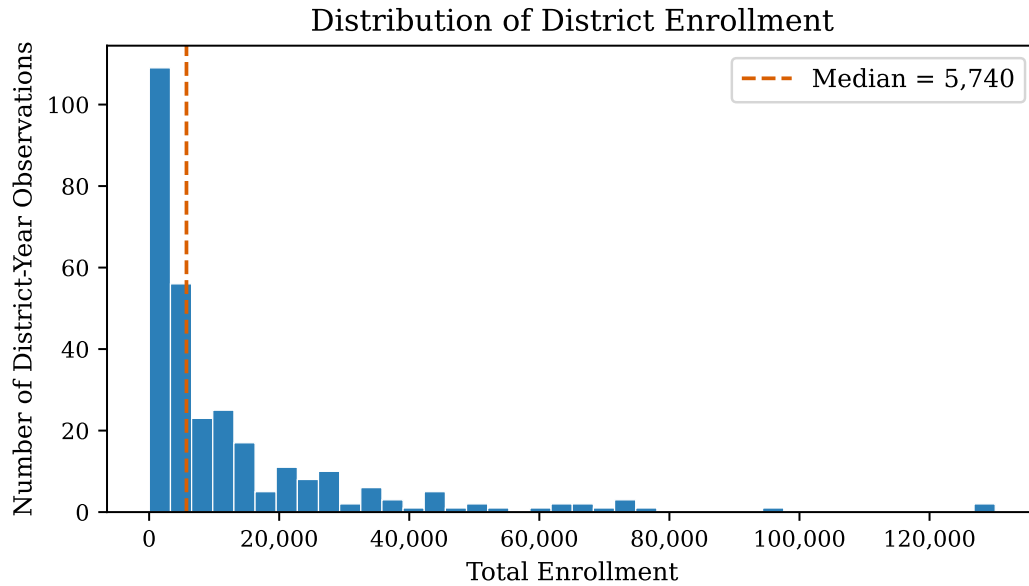
Figure 3: Distribution of district enrollment across all district-year observations.

Table 3: Party affiliation of candidates (where reported)

Party	Candidates
Nonpartisan (NON/NP)	913
Republican (REP)	114
Democrat (DEM)	102
Nonpartisan Undeclared (NONU)	37
Libertarian (LBT/LIB)	18
Green (GRN)	14
Independent (IND)	2

3.8   ADE Accountability Metrics

Table 4 reports the availability of each accountability metric across the 394 district-year observations.

Among the 167 district-year observations with letter grades, the distribution skews toward the higher end: 47 received an A, 62 a B, 51 a C, 5 a D, and 2 an F (Figure 8).

3.9   Temporal Coverage

Eight districts have data in all six election cycles: Mesa, Peoria, Scottsdale, Dysart, Cave Creek, Queen Creek, Deer Valley, and Kyrene. Table 5 shows the full distribution of year coverage across the 135 districts in the summary file.

In total, 104 of 135 districts (77%) appear in two or more election cycles, providing repeated observations for panel analysis.
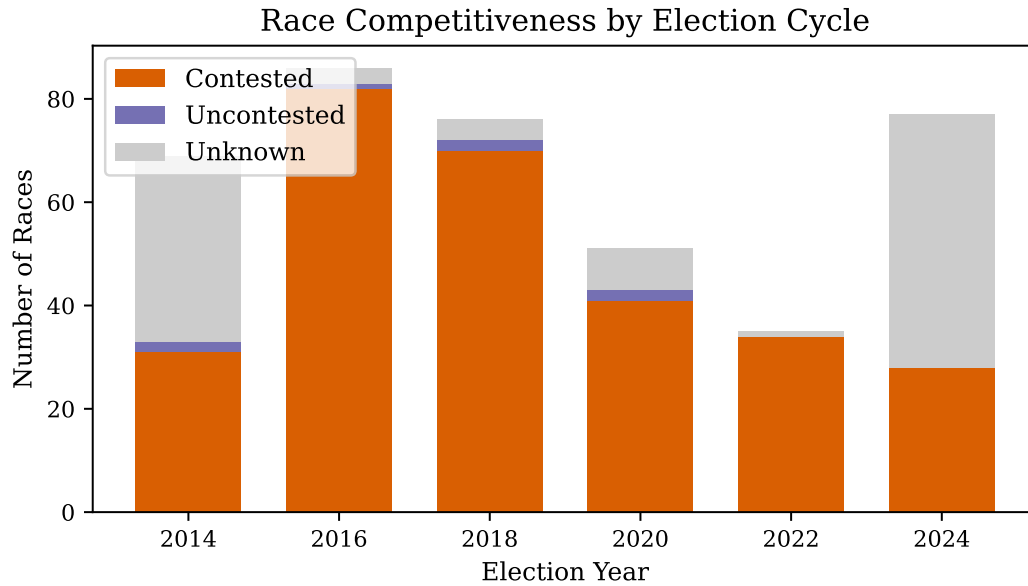
Race Competitiveness by Election Cycle

Figure 4: Number of contested, uncontested, and unknown races by election cycle.

Table 4: Accountability metric coverage

Metric	Available	Rate	Notes
A–F letter grade	167	42%	Not published in 2016 or 2020
Total enrollment	299	76%	
Dropout rate	273	69%	
4-year graduation rate	178	45%	Only districts with high schools
Superintendent	51	13%	Data collection in progress

4 Limitations and Known Issues

1. **Incomplete district coverage.** 85 of Arizona's 223 regular school districts are not yet in the dataset. Most are small rural districts, but several mid-size districts (e.g., Casa Grande Elementary, Prescott, Chinle, Snowflake) are also absent.
2. **Unbalanced panel.** Only 8 of 135 summary districts appear in all 6 cycles. The missing district-years are a mix of genuinely uncontested/off-cycle elections and data collection gaps.
3. **Seat counts for OpenElections records.** Races added from OpenElections precinct data often lack seat count information, making it impossible to determine whether the race was contested and to calculate candidates-per-seat ratios. This affects primarily the 2014 and 2024 additions.
4. **Winner identification.** The `winner` field is populated for only 587 of 2,984 candidate records (20%). This limits analysis of incumbent effects and winner characteristics.
5. **A–F grade gaps.** ADE did not publish school or district letter grades in FY2015–16 or FY2020–21, creating structural missingness in the 2016 and 2020 election cycles.
6. **Superintendent data.** Currently covers only 51 district-year observations (13%). Historical superintendent data back to 2013 is being collected.
7. **Party affiliation.** Available for 40% of records. Most Arizona school board races are officially
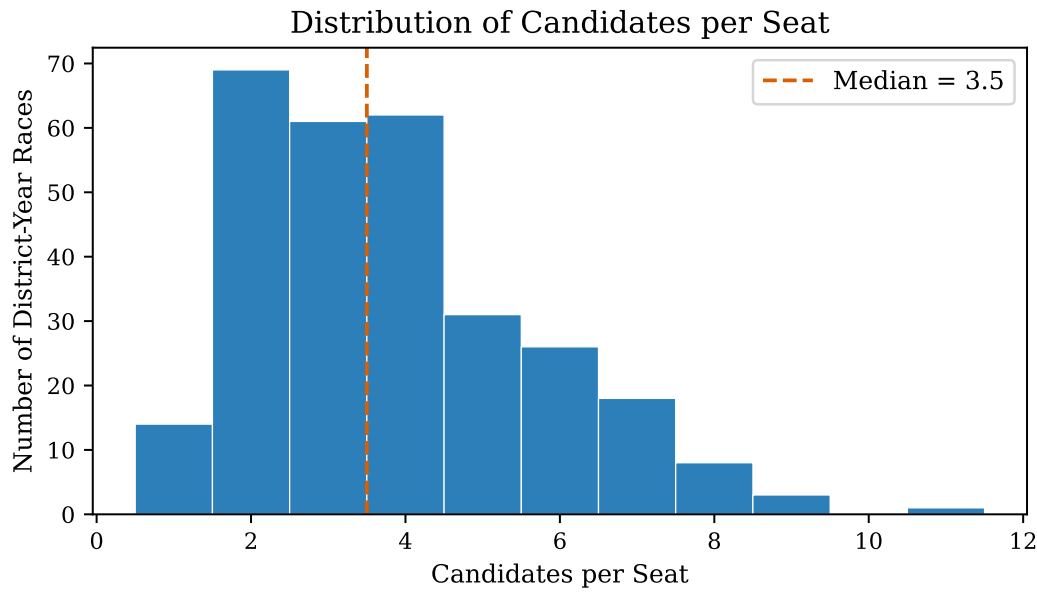
Distribution of Candidates per Seat

Figure 5: Distribution of candidates per seat across district-year races.

Table 5: Number of election cycles per district

Cycles Present	Districts	Cumulative %
6 (all years)	8	6%
5	15	17%
4	23	34%
3	26	53%
2	32	77%
1	31	100%
Total	**135**	

nonpartisan, so the party field reflects candidate filing data where reported by county election offices.

5 Data Dictionary Reference

Full column definitions for both the master and summary files are provided in `DATA_DICTIONARY.md` in the repository. Key variables include:

- `ctds_id` — ADE CTDS Entity ID (primary district identifier)
- `nces_leaid` — NCES national LEA identifier
- `year`, `election_date`, `election_type`
- `candidate`, `party`, `total_votes`, `winner`
- `contested`, `candidates_per_seat`, `winner_margin_pct`
- `lea_letter_grade`, `total_enrollment`, `dropout_rate`, `grad_rate_4yr`
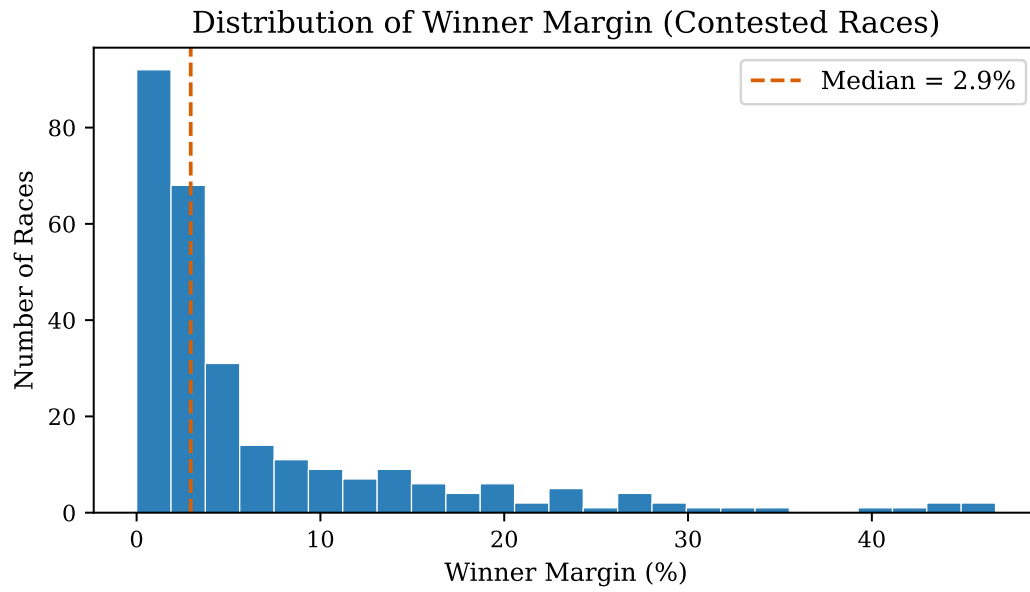
Figure 6: Distribution of winner margins in contested races ($\leq 50\%$).

Data files, source data, and ingestion scripts are all included in the repository.
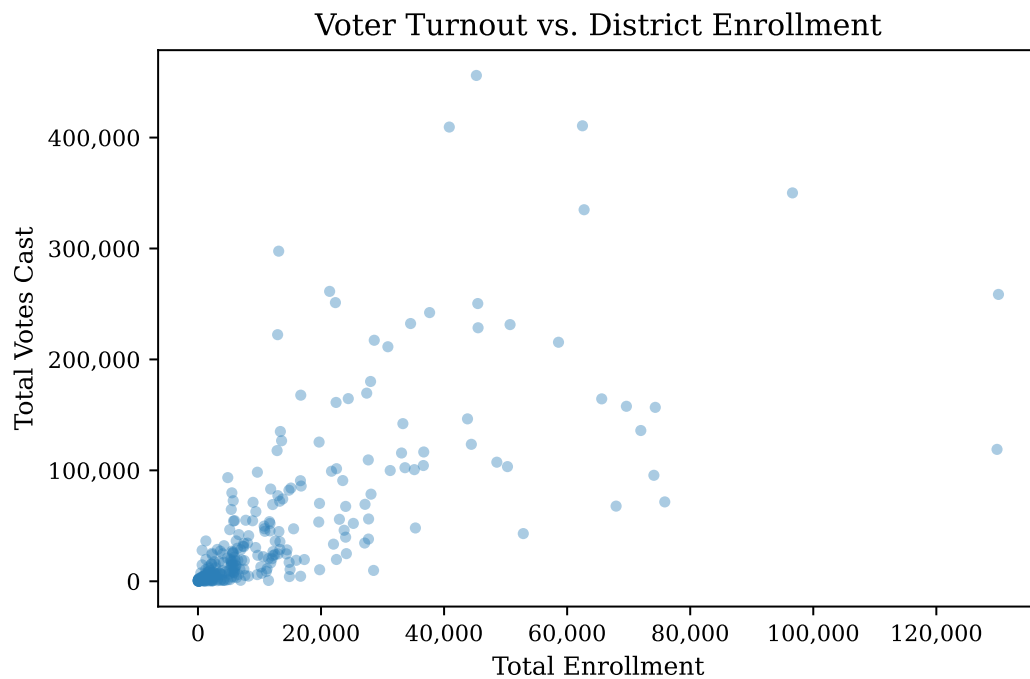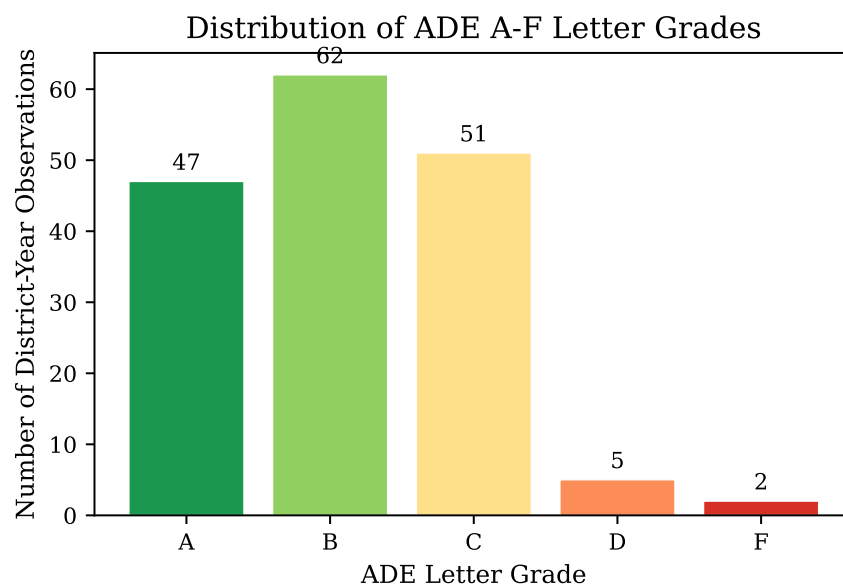
Figure 7: Total votes cast vs. district enrollment.

Figure 8: Distribution of ADE A–F letter grades.