Comparing the Comparators: Reviewing Reed College's Comparator List

Nat Lincoln, Acadia McDonald, and Leland Russell
Michael Pearce and Mike Tamada
Fall 2025

Abstract

Reed College maintains a 30-school comparator list used to benchmark institutional performance and planning, yet it is not always clear how closely these institutions continue to resemble Reed as colleges change over time. This study reassesses Reeds comparator group using a framework built from recent institutional data. Drawing on publicly available and internal data on student characteristics, academic outcomes, finances, and institutional resources, we compare Reed to a broad set of liberal arts colleges and evaluate its position within the existing comparator group. Institutions are compared using two complementary distance-based approaches, principal component analysis (PCA) and standardized Z-scores, to assess the stability of similarity patterns across modeling choices. Overall, Reed appears centrally located relative to its peers, suggesting broadly comparable positioning rather than consistent over- or under-performance on the dimensions. At the same time, several institutions emerge as consistently close comparators, while some commonly assumed peers appear less similar under the quantitative measures used. Based on these results, we propose two revised comparator lists of 24 and 26 institutions that are more focused and consistent than the current list. We also develop an interactive web application that allows users to adjust the relative importance of different attribute groups when generating comparator sets for a more specific set of comparators.

This project was for the fall 2025 term of Math 343: Statistics Practicum. It was given by Mike Tamada and advised by Professor Michael Pearce.

1. Introduction & Review

We start by reviewing existing literature relevant to the study of liberal arts colleges, college ranking systems, and the statistical methodologies commonly used to compare institutions of higher education. We will place liberal arts colleges within the broader higher education landscape, highlighting their defining characteristics, strengths, and ongoing challenges. We can then examine the development and use of college ranking systems, including their purposes and limitations. Finally, we discuss methodological approaches used in institutional comparison, with attention to weighting and interpretability. Together, these provide the conceptual and methodological foundation for the comparative framework developed in this study. With these in hand, we will end by discussing the goals of this project.

1.1. Liberal arts colleges.

Liberal arts colleges occupy a unique and historically significant position in the American higher education world. Originally founded to educate elite white men in the early nineteenth and twentieth centuries, these institutions emphasized broad intellectual development rather than narrow vocational training. As Hawkins (2000, as cited in [23]) explains, they were designed as four-year baccalaureate institutions resistant to highly specific vocational preparation and insisting on a considerable breadth of studies. This commitment to intellectual breadth rather than professional specialization remains the defining feature of liberal arts colleges today.

Liberal arts colleges are best known for their student-centered focus and emphasis on undergraduate teaching. With small enrollments, few, if any, graduate students, and strong faculty-student relationships, these institutions often prioritize mentorship, leadership development, and holistic education. Because liberal arts colleges devote nearly all their resources to educating undergraduates, rather than dividing attention among graduate programs, research, or professional schools, they are able to deliver more personalized instruction and foster close academic communities [9]. Many liberal arts colleges also pride themselves on producing future scholars. STEM majors, for example, often gain early laboratory experience that would typically be reserved for graduate students at large research universities. This contributes to strong graduation rates and to the disproportionate number of liberal arts alumni who later pursue Ph.D.s or return to teach in similar academic settings.

Despite their historical prestige and educational strengths, liberal arts colleges have faced growing financial and enrollment challenges over the past several decades. During the 2010s alone, over 20 liberal arts colleges closed due to financial instability, and another 40 merged with or were acquired by other institutions [23]. This trend is not new. Between 1972 and 1988, 334 liberal arts colleges either closed or reclassified themselves as different types of institutions [9]. Rising tuition costs have further intensified these pressures. In 2015, the average published price for attending a private nonprofit four-year college was \$43,921, compared to \$19,548 for an in-state public institution [23]. Even after accounting for financial aid, the net price of attendance remained nearly double that of public universities, and graduates from private colleges were more likely to carry student debt exceeding \$20,000 [23].

These financial challenges have been compounded by shifting attitudes toward higher education. Students and families increasingly prioritize career preparation and return on investment, making professional programs more attractive than traditional liberal arts curricula [11]. In response, many liberal arts colleges have attempted to adapt by adding pre-professional programs, de-emphasizing test scores in admissions, and expanding diversity initiatives [23]. Despite these efforts, Hilbun and Mamiseishvili describe liberal arts colleges as among the most financially and operationally vulnerable institutions in higher education [23]. Even as they adapt to modern demands, liberal arts colleges continue to share core values and missions that distinguish them from other higher education sectors. Most maintain a commitment to broad-based learning, small-scale education, and the development of critical thinking across disciplines. Breneman [9] cautions that while many colleges have survived by introducing professional or technical programs, doing so risks eroding the educational diversity that makes the liberal arts model valuable. He argues that the loss of distinctively liberal arts institutions would diminish the range of educational philosophies available to students in the U.S. system.

At the same time, liberal arts colleges differ in how they interpret and sustain these traditions. Some have become more specialized, integrating research, professional programs, or niche disciplinary focuses, while others adhere more closely to a classical liberal arts model. As Hu notes, although these colleges possess considerable freedom in shaping their missions, curricula, and admissions approaches, they all face the shared challenge of remaining true to a liberal arts identity while achieving financial sustainability and continued relevance [23].

1.2. Ranking systems.

College rankings in the United States emerged as a national concern during the late twentieth century, when the economy transitioned from production-based to knowledge-based industries. As higher education became a costly investment, consumers sought ways to quantify the quality of an institutions education [11].

Early college ranking systems were often conducted by institutions themselves and were criticized for being biased in favor of the institutions producing them [14]. Contemporary ranking systems are now developed by independent organizations such as *U.S. News and World Report*, *Times Higher Education World University Rankings*, and others [11], [5], [6], [7]. As rankings grew in popularity, so too did the number of stakeholders who relied on them. Institutional administrators, policymakers, and researchers use rankings to quantify institutional performance and to justify funding decisions [11], [14], [24].

Because ranking systems vary widely, so do the quality of indicators they employ. Commonly used variables include incoming class demographics, faculty-to-student ratios, and graduation rates [11], [24]. However, many of these indicators serve only as proxies and may not directly measure educational quality [11], [8]. Although quality is inherently subjective, Brasher et al. identify several recurring themes across ranking systems, including purposefulness in standards, distinctiveness, transformational impact, and accountability [8].

Ranking systems can be broadly divided into national and global frameworks. National systems typically use a greater number of indicators and place more emphasis on educational and institutional variables. Despite these differences, national and global rankings often produce similar results, suggesting that many of the underlying variables are correlated [24].

1.3. Statistical methodologies.

Methods for ranking institutions are typically based on weighting schemes, yet deriving meaning from these weights or from the resulting statistics is often difficult. Variables used in ranking systems are often dissimilar, making direct comparison challenging [15]. These difficulties are compounded by vague goals, poor variable measurement, and insufficient efforts to account for confounding factors [14], [8], [11]. Ranking systems also tend to undervalue non-numeric or non-ordinal data [24]. For example, school mission statements can provide important insight into institutional culture, yet are often excluded from rankings [8]. Moreover, these systems frequently treat all institutions as fundamentally comparable despite meaningful differences across institutional types [9], [22]. Higher research output, for instance, carries different implications for a research university than for a liberal arts college [8].

Despite these limitations, some approaches have proven more effective. The Scorecard or report card models group related attributes into broader categories, which are then weighted and compared [15], [6]. These models offer greater interpretability, as grouping similar variables simplifies interpretation and makes the overall framework easier to understand.

However, these models also introduce challenges. They require additional weighting decisions, and the weights assigned to individual variables within categories are often arbitrary [15]. To address this, some researchers propose statistical models that generate data-driven weights [15], [16]. While these approaches reduce human input, they often sacrifice interpretability and the method by which algorithms assign weights is not always well known. Consistency analyses are commonly used to interpret models by comparing outputs to established rankings [15], [16]. However, because established rankings themselves are imperfect, conclusions drawn from these comparisons may be limited, especially beyond inference about ranking systems. Regardless, consistency analyses to ensure results are reasonably stable are important.

1.4. Motivation.

Reed College maintains a 30-school comparator group that plays a central role in institutional assessment, planning, and comparison [4]. This list, most recently updated in 2022, is used by the Office of Institutional Research (OIR) and senior leadership to measure Reeds performance across a range of academic, financial, and operational dimensions. The comparator institutions were selected based on similarities in size, mission, finances, location, graduation rates, and related characteristics. The list is intended to provide a consistent frame of reference for evaluating Reeds standing relative to broadly similar liberal arts colleges.

The existence and use of comparator lists raise important questions about institutional similarity and usefulness. While comparator groups are often treated as reference sets, colleges evolve over time, and shifts in enrollment patterns, finances, and priorities can alter how comparable institutions truly are. Members of Reeds faculty and administration have noted that a comparator list of 30 institutions may be too large to be analytically precise, potentially obscuring meaningful differences and limiting its usefulness for focused comparison. These concerns motivate a careful reassessment of how Reed compares to its current reference group and whether the list continues to reflect institutions that are genuinely similar across dimensions. Our questions we wish to answer are as follows:

First, how does Reed compare to its existing 30-school group on updated measures, and are there indicators on which Reed consistently appears at the extreme high or low end of the distribution? Such situations are undesirable, as they may indicate that the comparison group no longer provides an informative benchmark.

Second, when looking at the broader scope of private liberal arts colleges, are there institutions that appear quantitatively similar to Reed but are not currently included, or conversely, institutions in the existing list that appear sufficiently dissimilar and may warrent reconsideration? Importantly, the OIR has emphasized the value of a similarity framework that is flexible and transparent, allowing users to adjust the relative importance of different attributes rather than relying on a single fixed weighting scheme.

Another focus of this project is the financial health of comparator institutions. Financial instability has increasingly affected liberal arts colleges, and institutions under severe financial pressure may no longer serve as appropriate benchmarks for Reeds long-term planning. However, this component was framed as a secondary, time-permitting extension of the core analysis rather than a primary objective of the project.

Together, these questions and motivations call for a systematic, data-driven approach to institutional comparison. This approach should reassesses Reeds existing comparator list, explore alternative sets of varying restrictiveness, and balance rigor with interpretability.

1.5. Relevance and goals.

This project aims to strengthen and extend this framework by systematically reviewing and updating institutional data, evaluating Reeds position relative to its comparison group, and developing a flexible method for identifying peer institutions. Specifically, we examine whether Reed has extreme values, either high or low, on key indicators relative to its peers. We also explore similarity measures that allow users to choose which attributes they value and the degree to which they value them.

To accomplish these goals, a model in which weights are determined entirely by statistical methods is undesirable, as is a model relying solely on user-generated weights. The former may introduce too much unknown, while the latter demands consistent user specification of weights for many variables. A middle ground may be scorecard-style frameworks, where variables are grouped into classes that are compared to one another. In this approach, model-selected weights for within-class variables can be combined with user-defined weights across classes. Techniques such as data-driven weights or principal component analysis may be useful in this context, as they allow groups of variables to be combined into single attributes that users can weight according to their preferences.

Finally, this study engages with two broader conceptual questions raised in the literature. First, it considers the ethical and practical implications of ranking institutions and whether comparative analyses reinforce hierarchical notions of educational quality. Second, it engages with ongoing concerns about the sustainability of liberal arts colleges by situating Reeds performance within a set of comparable liberal arts institutions, rather than evaluating sustainability outcomes directly. Ultimately, this project contributes to ongoing discussions about how liberal arts colleges can responsibly use comparative data to support their missions in a changing higher education landscape.

2. Data Acquisition & Description

Here we discuss our data sources, our methodology for collecting data from the sources, methods we used for preparing the data for analysis, and specific details for the variables used.

2.1. Data Collection. Our analysis began with a number of different datasets and data sources recommended by Mike Tamada of the OIR. Among these data sources were the following: the Integrated Postsecondary Data System (IPEDS) [19], the College Scorecard (CSC) tool [17], data from Federal Student Aid [18], National Science Foundation (NSF) data on doctoral degrees [21], data from the Voluntary Support of Education (VSE) [20], and Financial ratings from Moody's, Standard and Poor's, and Fitch [2], [3], [1]. There are more, but these are the main data sources we were considering. Of these, we chose the data from IPEDS, CSC, NSF, and VSE. We did not include the data on federal student aid or financial ratings because IPEDS had data on financial aid and obtaining financial data from the three firms was difficult.

The IPEDS is an annual, mandatory survey for all U.S. colleges that participate in federal financial aid programs under Title IV of the Higher Education Act. Each institution self-reports its data on enrollment, admissions, finances, and more. This was accessed directly from the IPEDS website [19].

The other three datasets, NSF, VSE, and CSC, were largely as provided to us by Mike Tamada, and we did not directly compile these data. The NSF Survey of Earned Doctorates is a freely available national dataset that collects information from everyone who earns a doctoral degree in the U.S. (a mandatory survey). The VSE data contains information about college donations and fundraising, which is self-reported by institutions on a voluntary basis and not publicly available. The College Scorecard data is a freely available national dataset that compiles institutional-level and field of study information. Datasets compiled in College Scorecard are IPEDS, National Student Loan Data System and U.S. Treasury data. This data is collected by the U.S. Department of Education and the U.S. Treasury and mandatory for institutions to report. All four sources rely on self-reported institutional data, but NSF, IPEDS, and College Scorecard are government-administered, while VSE is managed privately.

For the CSC, VSE, and NSF data, we used the data as was provided by Mike Tamada, preferring to keep most variables as is, but often chose averages over school size over total counts when possible. For the IPEDS data, we did an exhaustive search through the variables that are "Frequently Used" by IPEDS users. Although this initial selection may lead to selection bias, frequently used variables had a variety of themes such as institutional characteristics and graduation measures. That is, our selection was diverse in topic. Furthermore, there are a large number of variables in this group. It does, notably, exclude some important variables, religious affiliation to name one, so this is not a perfect method, although we chose to use it due to time constraints.

2.2. Initial Data Preparation. To clean the data from our four sources, we prioritized unique variables with recent and high responsiveness from schools in variable selection. Uniqueness often fell by the wayside, but when given a choice between counts variables and average variables, we opted to choose average variables and not use counts variables. This was so the measures in these variables didn't confound as strongly with school size.

For IPEDS, we used less than half of the original variables we selected from the site. These were for a number of reasons. When choosing whether to keep variables, we wanted them to be well represented in similar metrics. Diversity of admissions statistics, for example, were kept consistent within admissions statistics. So even though applications included data for other genders and unknown genders, since admission rates and admission yields did not include these variables, we did not include them for applications. Equally, many variables had very high missingness and thus were not suitable for use as it would require too much data imputation. As mentioned earlier, we calculated means across multiple years. When possible, we selected three school years (2021-2024) for every variable, but often had to select two school years (2021-2023) and averaged across them when they contained numeric information.

2.3. Data Imputation & Transformations. While we tried to keep variables with low missingness, there was still many schools lacking data for variables across the dataset. Furthermore, some variables which had large scales or particularly extreme values risked biasing models by making otherwise distant values seem closer together by contrast. To change the data via imputation or variable transformations, even with reasonable methods, raises a host of concerns. Data imputation may over or undervalue particular schools and variable transformations abstract meaning away from variables. Conversely though, leaving missingness and outliers as is may result in unpredictable behavior from models and under or overestimating the distance between schools. Despite the former concerns, we decided to impute missing variables and apply variable transformations.

Variable transformations were only done with logarithms. We decided whether or not to use a log transform on a school depending on if the quantiles appeared less skewed under a log transformation. There was not a hard or fast rule for determining which variables got transformations, though we only used it on a few variables. These variables were largely related to donations and enrollments—counts variables which naturally could be very skewed based on school size.

We had two methods for imputing data. These were with means and medians. These resulted in similar but slightly different results. For median data imputation, we simply replaced any missing value of a variable with the median of that variable.

This was without regard to variable transformation as the median is invariant under the logarithm. For means, however, since the logarithm of the mean is distinct from the mean of the logarithm, the choice of imputing the raw data or transformed data must be intentional. The former gives values which are more representative to the real observed values, however it is equally true that we are taking log transformations to decrease skewness, and the means taken before logarithms would retain some of this skewness. Taking the mean of the logarithm would remove some of the effect of these outliers and give a more robust statistic. As such, we imputed the data with means after taking variable transformations.

2.4. Variable Grouping. To make these variables more suitable for analysis, we developed a grouping framework that would allow us to summarize related variables using Principal Component Analysis (PCA) and Z-scores. Without grouping, we would have needed to generate an unwieldy number of summary statistics, making distance calculations between schools difficult to interpret, and harder to modify in a consistent way. Grouping enabled us to retain the breadth of the data while reducing dimensionality. Furthermore, for methods like PCA, particular principal components lose much meaning. Likewise when computing a distance from Reed within groups for Z-scores, interpretation becomes nearly impossible. These groups reintroduce some meaning to these otherwise unwiedly statistics, allowing us to make meaningful conclusions based on our results.

We selected eight groups intentionally, preferring a larger number of coherent, narrowly defined categories over a smaller number of overly broad ones. Fewer groups would have risked blending variables that reflect distinct concepts, reducing interpretability and possibly biasing PCA results. At the same time, we aimed to keep the groups reasonably balanced in size so each group was roughly equally representative of the data. The grouping process drew on standard classification frameworks used in college ranking systems [**11**], [**24**], recommendations from Mike Tamada, and our own judgment regarding which variables naturally belonged together.

The final structure consisted of eight conceptually cohesive groups: Applicant Demographics, Applicant Test Scores (often just called Test Scores), Alumni Donations, Enrolled Student Demographics, Graduation Outcomes, Staff and Faculty, Institution Finances, and Cost for Students. For example, all SAT and ACT percentile measures were placed together in the Test Scores group, while all variables describing financial burden or aid eligibility were included in the Cost for Students group. Likewise, variables capturing retention, graduation rates, or postgraduate earnings formed the Graduation Outcomes group, and those relating to institutional spending, endowment, and financial ratios constituted the Institution Finances group. Across the remaining groups, the same logic was applied to ensure that variables captured the same underlying idea and that the groups were not overly focused but not overly vague. There is a codebook describing the variables and their groupings in Table 1 and Table 2.

This grouping system provided the structural foundation for subsequent PCA and Z-score-based analyses. By organizing variables into clear and balanced categories, we were able to derive summary measures that preserved the complexity of the original dataset while allowing for transparent comparisons across institutions.

Grouping	Variables
Applicant Demographics	Total Applicants; Total women applicants; Total men applicants; Percent of admitted students; Percent of admitted women; Percent of admitted men; Total Admissions yield; Admissions yield for women; Admissions yield for men
Applicant Test Scores	SAT math 25th, 50th, and 75th percentiles; SAT reading and writing 25th, 50th, and 75th percentiles; ACT composite 25th, 50th, and 75th percentiles
Alumni Donations	Average alumni donors; Average alumni solicited; Average total alumni donation; Average enrollment; Average percent alumni donors; average donation per alum; Average donation per student; Average gift 2022; Average gift per student
Enrolled Student Demographics	Enrollment; Percent of students American or Alaskan native; Percent of students Asian, native Hawaiian, or pacific islander; Percent of students African American; Percent of students women; Historically black college or university
Graduation Outcomes	Total graduation rate; Retention rate; Graduation rate 6y women; Graduation rate 6y men; Graduation rate 6y pell grant; Graduation rate 6y no loans; Percent doctorates; median earn working not enrolled graudates
Staff and Faculty	Total salaries and wages; Total staff; Instruction research FTE; library services FTE; education services FTE; management FTE; computer, engineering, and science FTE; service, legal, arts, and media FTE; healthcare FTE
Institution Finances	Instruction expenses; Endowment assets per student; Student service expenses; Endowment assets; Equity Ratio; Library expenditures per FTE; Salaries as percent of core expenses
Cost for Students	Total out of state price; Total in state price; Percent of students with financial aid; Average price for students on grant

Table 1. Groupings for all variables. Semicolons separate variables, but for test scores, multiple are listed at once (SAT 25th, 50th, and 75th percentile represents three variables).

Variable	Description
ACT Composite	25th, 50th, and 75th percentile of reported ACT scores in applicants
Admissions Yield	Percent of students admitted who accepted the offer. Demographic information present for women and men
Applicants	Number of students who applied. Demographic information present for women and men

Average Price for Students on Grant	The average price for students on federal or institutional grants.
Average Alumni Donors	Average number of alumni who donated.
Average Alumni Solicited	Average number of alumni solicited for donations.
Average Enrollment	Average number of undergraduate students
Average Gift per Student 2022	Average nominal deferred gift for the year 2022 per student
Average Donation per Alum	Average giving per alumni that donates
Average Donation per Student	Average giving per student
Average Gift 2022	Average nominal deferred gift for the year 2022
Average Percent Alumni Donors	Average percent of alumni that donate after being solicited
Average Total Alumni Donation	Average total giving from the alumni that donated
Computer, Engineering, Science FTE	Total number of full-time computer, engineering, and science staff
Education Services FTE	Student and academic affairs full time equivalent staff
Endowment Assets per Student	Year end endowment assets divided by the full time equivalent student enrollment
Enrollment	Total number of undergraduate students enrolled for credit
Equity Ratio	Total net assets divided by total assets
Graduation Rate 6y Men	Graduation rate of men within the full time, first time students within six years
Graduation Rate 6y No Loans	Graduation rate of students with no loans or aid within the full time, first time students within six years
Graduation Rate 6y Pell Grant	Graduation rate of students eligible for a pell grant within the full time, first time students within six years
Graduation Rate 6y Women	Graduation rate of women within the full time, first time students within six years
Graduation Rate Total	Graduation rate of total cohort within six years
Healthcare FTE	Total number of full-time Healthcare staff
Historically Black College or University	Recognized status as a Historically Black College or University
Instruction Expenses	Instruction expenses as a percentage of core expenses
Instructional Research FTE	Full time equivalent counts of staff for instructional, research, and public service staff
Library Expenditures per FTE	Total library expenditures divided by full time equivalent students
Library Services FTE	Full time equivalent staff for librarians, curators, and other teaching instruction support staff
Management FTE	Full time equivalent management staff

Median Earn working not enrolled graduates	Median income of graduates working a job and not enrolled in an educational institution. Taken 10 years post graduation
Percent Doctorates	Average percent of alumni that earned a Doctoral Degree
Percent Admitted Men	Percent of men admitted to a institution
Percent Admitted Total	Percent of applicants admitted to a institution
Percent Admitted Women	Percent of women admitted to a institution
Percent of Students with Financial Aid	Percent of students receiving any financial aid or scholarship funding
Percent Students African American	Percent of students that are African American
Percent Students American Alaska Native	Percent of students that are American or Alaskan Native
Percent Students Asian, Native Hawaiian, or Pacific Islander	Percent of students that are Asian, Native Hawaiian, or Pacific Islander
Percent Students Women	Percent of students that are women
Retention Rate	Full-time retention rate
Salaries as Percent of Core Expenses	Salaries and wages as a percentage of core expenses
SAT Math	25th, 50th, and 75th percentile of reported SAT math scores in applicants
SAT Read Write	25th, 50th, and 75th percentile of reported SAT math scores in applicants
Service Legal Arts Media FTE	Total number of Community Service, Legal, Arts, and Media staff
Student Service Expenses	Student service expenses as a percent of the total core expenses per year.
Total In-State Price	Total tutition price for in-state students living on-campus
Total Out-of-State Price	Total tutition price for out-of-state students living on-campus
Total Salaries and Wages	Total salaries and wages as a percentage of total expenses
Total Staff	Total number of full-time staff at a institution

Table 2. Qualitative descriptions for variables used. A technical codebook has been created in Appendix A containing variable names used in the code and more precise descriptions. The groupings of variables are in Table 1. There are also two nearly repeated variables: average enrollment, from alumni donations, and total enrollment, from student demographics. These are both important variables for contextualizing other attributes within their respective gouprs and hence included.

3. Methods

To analyze the grouped data, we used two methods: PCA and Z-Scores. PCA was chosen as a way to group variables within groups, allowing for just one or two summary statistics to represent an entire group of data. Z-Scores were chosen as an alternative method for compressing the data. Each of these results in some loss of data, although the resulting statistics are much more meaningful and streamlines the specification of weights and later interpretation of results.

3.1. PCA Distance Methods. PCA has the advantage of summarizing variables by leveraging attributes like correlation to find statistics which capture as much information as possible. It is a widely used technique for data compression and dimension reduction [12].

When using PCA, with our data, we chose to retain enough principal components to explain 65% of the variance of the group data. In this case, we had between 1 and 3 principal components per group. This resulted in eight vectors representing distance, one for each group. From here, we computed the group-wise euclidean distance from each school to Reed, summing across the principal components. The vector of group distances for a school can thus be represented by the following:

$$\begin{pmatrix} \sqrt{\sum_{j \in \#\text{Principal Components}} \left(X^{\text{PC } j}_{\text{Reed, group 1}} - \vec{X}^{\text{PC } j}_{\text{group 1}} \right)^2} \\ \vdots \\ \sqrt{\sum_{j \in \#\text{Principal Components}} \left(X^{\text{PC } j}_{\text{Reed, group 8}} - \vec{X}^{\text{PC } j}_{\text{group 8}} \right)^2} \end{pmatrix}.$$

This vector could still be subject to bias though. For groups with many principal components, their distances may be substantially higher on account of including more terms. To ensure that this was not an issue, we then standardized the variance of each group distance to be one.

3.2. Z-score Distance Methods. For Z-scores, we used standard Z-score computations as an alternative to PCA. This gives an indication as to how consistent our later methods are, as we have very similar techniques between PCA and Z-scores. It is also consistent with prior OIR analyses of similar data. Mike Tamada has said that he typically uses Z-scores for similar analyses, so this will be consistent with those to some degree. A standard Z-score is calculated as

$$z = \frac{x - \mu}{\sigma},$$

where x is the observed value, μ is the mean of the variable, and σ is the standard deviation of the variable. The Z-score is the number of standard deviations a value is from the mean. Z-scores don't require that variables are normally distributed. However, they are more informative if the variable's distribution is more symmetric.

After calculating Z-scores for variables, as we did with PCA, we calculated the standard Euclidean distances from Reed's position for our eight groups. That is, we got the distance vector:

$$\begin{pmatrix} \sqrt{\sum_{j \in \#\text{Z-score}} \left(X^{\text{Z-score } j}_{\text{Reed, group 1}} - \vec{X}^{\text{Z-score } j}_{\text{group 1}} \right)^2} \\ \vdots \\ \sqrt{\sum_{j \in \#\text{Z-score}} \left(X^{\text{Z-score } j}_{\text{Reed, group 8}} - \vec{X}^{\text{Z-score } j}_{\text{group 8}} \right)^2} \end{pmatrix}.$$

Again, like with PCA, we normalized the variance of each group distance to be one to account for variable group sizes.

3.3. Differences in Methods. As briefly mentioned before, our two methods have distinct features. PCA produces only a few summary statistics for a group. This significantly reduces the dimensionality for further analysis. On the other hand, Z-scores preserves the dimensionality of the groups as each Z-score is a summary of a single variable. This is not necessarily an issue since we have normalized the variance of group distances.

We used PCA as an alternative to Z-scores as it is able to leverage variable associations to create statistics that explain the whole data in a more coherent way. So two highly correlated variables can be explained by one variable containing most of the signal in the two, and then a residual second variable with the remaining variability. Likewise, with many variables, it may be able to use elements of all of them to create just a few summaries.

Ultimately, Z-scores and PCA are very similar methods though, and we have applied them in quite narrow situations—there are not a lot of variables per group. So the two methods, both summarizing data, will likely not differ much. Because they are similar, though, we can evaluate the consistency of later methods using the differences between these two. Likewise, using mean data imputation versus median data imputation provides another very similar but somewhat distinct way of summarizing the data.

3.4. Application of groups. Once we have the groups, we would like to create a list of schools. One way to naively do this is to simply take the euclidean norm of each school's eight group distances to get a single distance value. So, we compute that the distance is

$$d(\text{School, Reed}) = \sqrt{\sum_{i=1}^{8} (\text{Group Distance}_i)^2}.$$

Since the group distances already represent distances from Reed, simply taking the norm again represents a distance from Reed. Then, sorting these, we have a list of schools representing some sort of closeness to Reed. This may be a good way to get lists of close schools to Reed, and such lists are proposed in Section 4.3. However, there are many ways to get candidate comparators with this list of group distances.

We propose a variable weighting method. It requires a user inputted weight c giving a numerical preference to each group, and multiplying each coordinate by such the corresponding value. So, with weight vector c,

$$d(\text{School,Reed; } c) = \sqrt{\sum_{i=1}^{8} (c_i \cdot \text{Group Distance}_i)^2}.$$

This can of course be transformed into the previous method by choosing $c = (1, \ldots, 1)$. The benefit of this, though, is that it permits a user to create specialized candidate comparator list. So an administrator interested in alumni donations and graduate outcomes, but not with admission statistics may weight the former groups higher and eliminate the latter groups by assigning a weight of zero. Conversely, an admissions worker may be interested in what schools are close to Reed in admissions statistics and campus culture. This method allows one to create a taylored list for a particular use.

4. Results

We will discuss how our current methods can answer the question on how Reed compares to the current comparator list. Then, we will compare the different models and create proposal comparator lists through the equally weighted method described in section 3.4. Finally, we will discuss a web-app framework to allow other users to interact with our analyses.

4.1. Testing Reed's Centrality. One of the main goals of this project was to determine if where Reed falls relative to current comparators. Using the biplots from PCA in Figure 1, we see that Reed is reasonably close to the center of principal components. Of course, this does not necessarily mean that it is for certain central in each attribute—PCA creates very hard to interpret statistics. The fact that it is not at extremes is promising though.

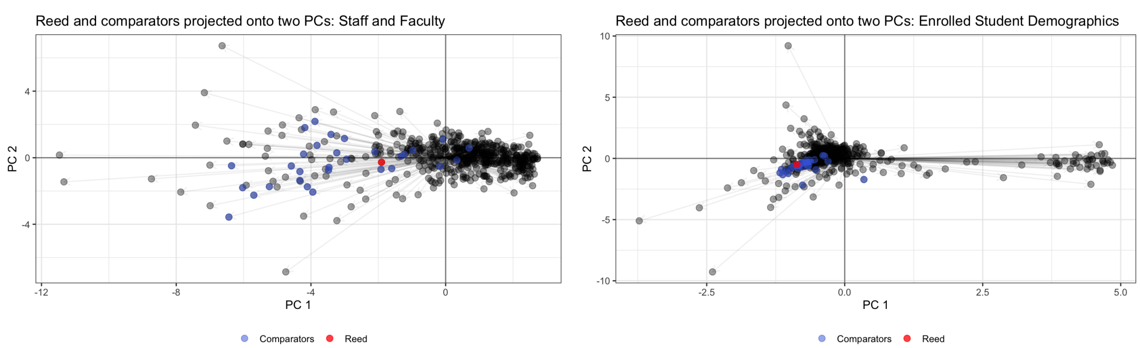

FIGURE 1. Two-dimensional biplots of principal components for staff and faculty (left) and enrolled student demographics (right) with mean data imputation. Reed is in red and comparators are blue. The remaining biplots are in Appendix B.

This is only shown for PCA means. We have left out PCA medians as the plots are not significantly different. The remaining PCA means plots, though, are in Appendix B.

This follows for Z-scores too. In Figure 2, we see that for all groups, Reed is surrounded by a cluster of other schools. This implies that Reed isn't red-lining or

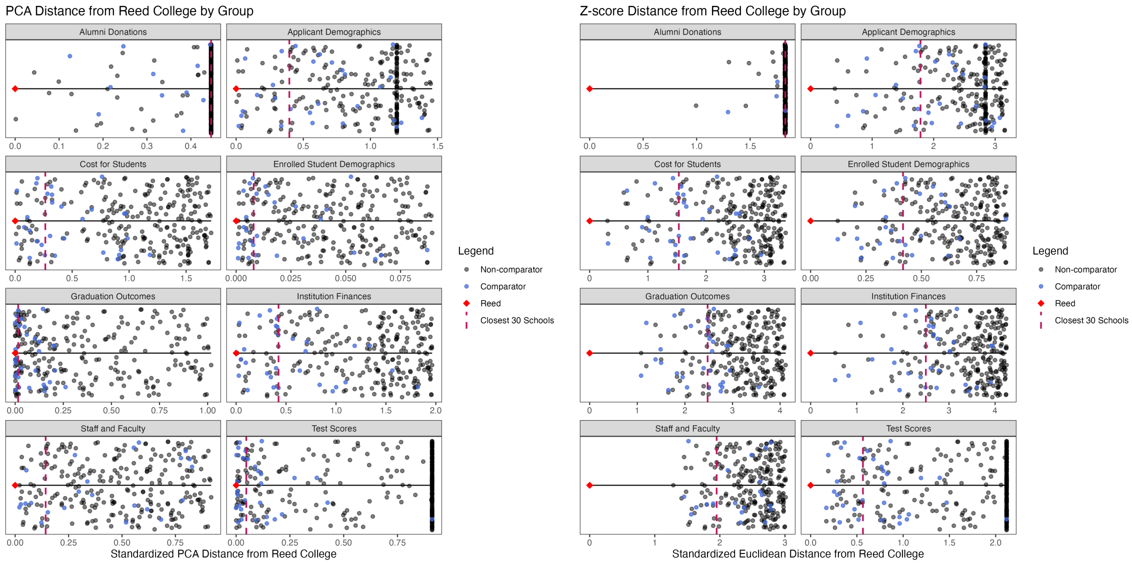

FIGURE 2. This graph shows how close any school less than the median value of a group is to Reed College for each variable grouping. The red diamond represents Reed College, the blue dots represents schools that are in the current comparator list, and black dots represent schools that are not in the current comparator list. The pink line represents the distance of the 30th closest school to Reed College. `NA` values were imputed with the mean of a variable.

green-lining as there are many schools close in distance. For example, with Graduation Outcomes, the line indicating the closest 30 schools is almost the value of Reed. Another example is in Enrolled Student Demographics. The line indicating the closest 30 schools is very close to Reed. Because some variables groupings had significantly more missing data, the total number of schools less than the median is not 250 schools as expected. Groups that were affected are Alumni Donations and Test Scores.

This is only shown for mean imputed methods. We have left out median imputed methods as the plots are not significantly different. The remaining plots, though, are in Appendix C.

4.2. Model Comparison. Across all four analytical approaches—Z-score imputed means, Z-score imputed medians, PCA imputed means, and PCA imputed medians—several consistent patterns emerged. Most strikingly, Whitman College and Haverford college ranked either closest or second closest in every method, suggesting a robust similarity to Reed under all versions of our distance metrics. This convergence across methods increases confidence that Whitmans and Haverford's positions are not an artifact of a particular statistical choice, but reflects a genuinely close institutional profile relative to the dimensions we analyzed.

A more unexpected result concerns the University of the South (Sewanee), which appeared within the top tier of comparators in all four methods. Culturally and institutionally, Sewanee is typically seen as quite different from Reed, and it has not historically figured prominently in comparator discussions. Its repeated appearance raises two possibilities. First, it may indicate that Sewanee aligns more closely with Reed on the quantitative dimensions we selected than previously assumed. Alternatively, this result may suggest there are features or weighting decisions in our methodology that warrant closer examination. In particular, factors that Reed prioritizes but that are not well captured by our grouped variables may contribute to this unexpected similarity.

Conversely, Lewis & Clark, commonly recognized as one of Reeds closest peers and included in the original comparator list, ranked unusually low in all four methods. This pattern may imply that Lewis & Clark is less similar to Reed across the specific quantitative dimensions we measured than commonly held assumptions suggests. On the other hand, it may indicate that our current grouping or weighting structure does not fully capture the attributes that make the two institutions comparable. This discrepancy highlights the need to reflect on whether there are important qualitative or institutional characteristics absent from our numerical analysis.

When examining the broader set of schools, we found that 11 institutions appeared across all four methods and were also part of Reeds original comparator list, aligning with prior expectations and providing face validity for the approach. In addition, 13 institutions appeared across all methods despite not being included in the original list. This suggests that our methods are identifying plausible comparators that may have been previously overlooked. These counts are based on restricting attention to the top 30 institutions in each method, consistent with the size of Reeds existing comparator set and our goal to avoid expanding the list beyond the scale typically used. The significant overlap also confirms, in part, the consistency between the methods. 24 out of the top 30 schools were shared between the four, and another two were shared between three. This is a high degree of similarity and confirms that the analysis is quite consistent with slightly different data.

To actually compare the models though, we have constructed distance plots, shown in Figure 3.The distance scales between PCA and Z-scores are different but that is inconsequential and a product of different methods for summarizing the data. Instead, we see that having groups which are significantly further from Reed in any coordinate is rarer in PCA than Z-scores. While this trend lessens as we get to schools further from Reed, the schools closer to Reed by PCA metrics are uniformly closer across most of their coordinates. The schools and distances generated with Z-scores, however, often have multiple coordinates which are significantly further from Reed. This does not mean that the PCA methods are necessarily a better method for getting close schools per se. It means that the lists generated with the PCA generally have less variability among the distance coordinates. This is somewhat expected as PCA can take advantage of correlated variables to reduce redundancies in summary statistics. It is certainly desirable. However, considering that the two methods came to very similar lists, we prefer the PCA list, but recognize that Z-scores is an equally valid list.

4.3. Proposed new lists. Building on the model comparison results, we propose two updated comparator lists derived directly from the consistency of institutions appearances across the four analytical approaches. These lists are intended to complement, rather than replace outright, Reeds existing 30-school comparator set by offering empirically grounded alternatives with different levels of restrictiveness.

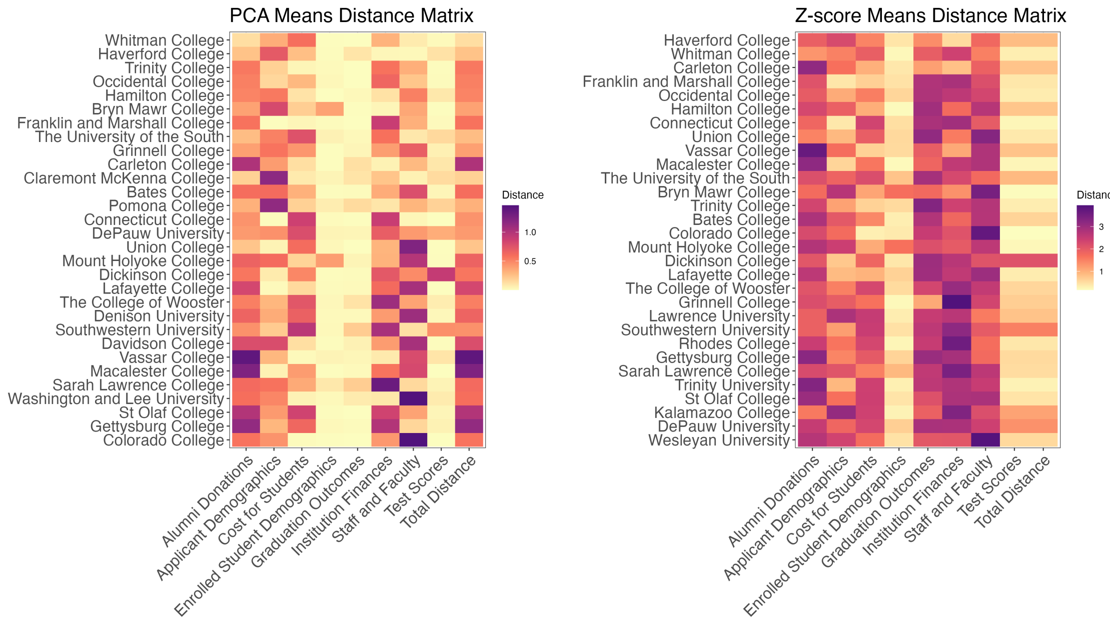

FIGURE 3. Distance matrices for PCA and Z-scores, both with Data imputation. The x-axis shows groups and the y-axis shows schools. Color scales range from zero to the maximum value. These only show mean data imputation. Median data imputation is in Appendix D

The first proposed list is a more conservative revision, consisting of 24 institutions, shown in Table 3. These schools appeared within the top 30 closest institutions to Reed under all four methods—Z-score imputed means, Z-score imputed medians, PCA imputed means, and PCA imputed medians—indicating strong robustness to modeling choices.

Bates College	Macalester College
Bryn Mawr College	Mount Holyoke College
Carleton College	Occidental College
Colorado College	Sarah Lawrence College
Connecticut College	Southwestern University
DePauw University	St. Olaf College
Dickinson College	The College of Wooster
Franklin and Marshall College	The University of the South
Grinnell College	Trinity College
Hamilton College	Union College
Haverford College	Vassar College
Lafayette College	Whitman College

TABLE 3. Proposed Comparator List (24 Schools)

The second proposed list expands this framework to include 26 institutions, this is in Table 4. In addition to the 24 schools described above, this list incorporates two additional institutions that appeared in the top 30 under two or three of the four methods. This broader list reflects a more inclusive approach, allowing for some methodological sensitivity while still grounding inclusion in quantitative similarity.

Bates College	Macalester College
Bryn Mawr College	Mount Holyoke College
Carleton College	Occidental College
Colorado College	**Rhodes College**
Connecticut College	Sarah Lawrence College
DePauw University	Southwestern University
Dickinson College	St. Olaf College
Franklin and Marshall College	The College of Wooster
Gettysburg College	The University of the South
Grinnell College	Trinity College
Hamilton College	Union College
Haverford College	Vassar College
Lafayette College	Whitman College

Table 4. Proposed Comparator List (26 Schools). The two added schools, Gettysburg College and Rhodes College, are bolded.

Relative to Reeds original 30-school comparator list, these proposed lists represent a notable shift. Several institutions from the original comparator set do not appear in either of the new lists. Several schools commonly assumed to be close peers, such as Lewis & Clark College, Willamette University, and Pomona College, were expected to appear but ranked substantially farther from Reed under the quantitative measures used. Other institutions, including the University of Puget Sound and Wesleyan University, were close to inclusion but fell just outside the cutoff thresholds. In contrast, some schools that were not historically emphasized in comparator discussions emerged as consistently similar under multiple methods. In addition, we intentionally sought to keep both proposed lists below 30 institutions, reflecting prior feedback that Reeds existing 30-school comparator set may be too large to be useful.

These exclusions do not imply that removed institutions are inappropriate comparators in all contexts. Rather, they suggest that across the specific quantitative dimensions analyzed here, some schools are less similar to Reed than conventional wisdom might suggest. The proposed comparator lists also retain meaningful continuity with Reeds original 30-school comparator set. 12 institutions from the original list appear in both the more restrictive 24-school list and the expanded 26-school list. This degree of overlap suggests that the quantitative methods used here recover a substantial portion of Reeds historically identified peers, providing face validity for the approach. At the same time, differences between the original and proposed lists highlight how systematic, data-driven methods can both confirm established assumptions and surface.

4.4. Interactive App. To accompany these models, we have created an interactive web-app through the Shiny framework [10], [13]. This allows a user to select a specific model we created, set a preference on group importance, then create a possible comparator list. This has the benefit of providing lists of close schools with equal weights, as we have used thus far, as well as more dynamic weights. This can create a specialized list of comparator colleges for a particular decision.

In addition to dynamic weighting, new methods may be added to the app. The app searches for all csv files in its current folder and can apply the analysis to all. If there is a another analytical tool, its data may be incorporated into the app provided it is in a standard form. Such a method must work within the existing eight group structure implemented here and return a single value for each school-group observation. Further details are discussed in Appendix E.

A limitation of the app is that users will only know if a school is close to Reed, not if that school outperforms or underperforms Reed. For example, if the closest school is Whitman College based on a user's preference, there is no information on whether Whitman is "better" than Reed in all groups. Because the app does not display this information, it leaves users unsure of if Reed red-lines or green-lines based on their preferences.

5. Discussion

This section discusses the results of the comparator analysis, placing the findings within broader methodological and institutional context. It highlights key patterns in the results, addresses limitations of the approach, and outlines implications for future institutional research.

5.1. Interpretation of results. Across multiple modeling choices, the analysis produced a set of comparator institutions that was both internally consistent and partially aligned with Reeds existing comparator list. The stability of certain institutions, most notably Whitman and Haverford College, across all analytical approaches suggests that the framework captures meaningful dimensions of institutional similarity rather than artifacts of specific methodological choices. At the same time, variation in the inclusion of other schools underscores the sensitivity of comparator identification to variable selection and grouping decisions. Across the variables examined, Reed did not consistently appear at extreme values relative to its comparators. Rather than exhibiting systematic red lining or green lining on particular dimensions, Reeds profile tended to fall within the central range of the comparator distributions, suggesting broadly comparable positioning across the measured indicators.

The results also indicate that several institutions commonly assumed to be close peers did not emerge as strong comparators under the methods used. This outcome does not imply that these institutions are inappropriate points of reference in all contexts. Rather, it suggests that the selected variables and grouping structure may not fully capture certain dimensions of similarity that are meaningful in practice. In particular, the limited representation of Pacific Northwest peers suggests that regional context, shared recruitment markets, or cultural proximity may not be adequately reflected in the current model.

The variability in distances between the PCA and Z-score methods are also interesting. While offering some support for the usage of PCA, it is important to recognize that the two methods came to similar lists. Further, the Z-scores methods treat each variable as unique from others in its group. This is not ideal, but explains the sometimes high distances from Reed in certain coordinates present in the Z-scores data. We thus recommend PCA over Z-scores but with the understanding that both methods should be used in conjunction with each other.

It is worth considering as we examine the results that we have chosen a particular set of variables that we thought would work a priori. The danger of selection bias is certainly present. By appealing to what we thought was reasonable and the advice of Mike Tamada, we hope to minimize this risk, but it is nonetheless still present. With any analysis of this subject, though, it will be impossible to fully remove. It is hence important to consider other possibilities for choices like variable grouping, variable selection, variable transformations, data imputation, and similar choices, and the alternatives possible.

Finally, it is important to note that one of the projects original stretch goals, developing a standalone predictive model of institutional financial distress, was not fully realized within the scope of this study. While financial variables were incorporated into the broader framework, constructing and validating a model capable of reliably predicting financial instability among liberal arts colleges would require additional data, and given the time constraints and scope of the project, the priority was placed on comparator identification. The financial model goal was treated as an exploratory extension rather than a core deliverable. Nonetheless, the preliminary considerations outlined here suggest that such a model is feasible and represent a direction for future work.

5.2. Methodological considerations and limitations. This study highlights the extent to which substantial preprocessing is required when conducting quantitative comparisons across higher education institutions. Many variables required log transformations and standardization to ensure that differences in scale did not disproportionately influence distance calculations. While these steps are standard in multivariate analysis, they introduce additional modeling decisions that can affect results. The need for heavy preprocessing reflects the inherent difficulty of comparing institutions that differ in size, resources, and organizational structure, even within the relatively homogeneous category of liberal arts colleges.

In addition, while financial variables were included among the grouped indicators, this analysis does not constitute a comprehensive model of institutional financial health. Similarity in this framework should therefore be interpreted as descriptive rather than predictive. Institutions that appear quantitatively similar at a given point in time may face very different long-term financial risks or opportunities, a distinction that is not captured by the present approach.

References

[1] Fitch Ratings: Credit Ratings & Analysis For Financial Markets.

[2] Moody's - credit ratings, research, and data for global capital markets.

[3] S&P Global Ratings.

[4] Comparison Schools for Reed College, 2022.

[5] Academic Ranking of World Universities. Technical report, Shanghai Ranking, 2025.

[6] Time Higher Education World University Ranking. Technical report, Times Higher Education, 2025.

[7] US News and World Report. Technical report, 2025.

[8] Andrew Brasher, Denise Whitelock, Wayne Holmes, Francesca Pozzi, Donatella Persico, Flavio Manganello, Marcello Passarelli, and Albert Sangrà. Comparing the comparators: How should the quality of education offered by online universities be evaluated? *European Journal of Education*, 57(2):306–324, June 2022.

[9] David W. Breneman. *Liberal Arts Colleges: Thriving, Surviving, or Endangered?* Brookings Institution Press, Washington, 1994.

[10] Winston Chang, Joe Cheng, JJ Allaire, Carson Sievert, Barret Schloerke, Yihui Xie, Jeff Allen, Jonathan McPherson, Alan Dipert, and Barbara Borges. *shiny: Web Application Framework for R*, 2025. R package version 1.11.1.

[11] Meredith Davis. Can College Rankings Be Believed? *She Ji: The Journal of Design, Economics, and Innovation*, 2(3):215–230, 2016.

[12] Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, and Jonathan E. Taylor. *An introduction to statistical learning: with applications in Python*. Springer texts in statistics. Springer, Cham, 2023.

[13] Leland-Russell. Leland-Russell/Math343_comparator_colleges_public, December 2025. original-date: 2025-12-17T00:59:12Z.

[14] Marc Meredith. Why Do Universities Compete in the Ratings Game? An Empirical Analysis of the Effects of the U.S. News and World Report College Rankings. *Research in Higher Education*, 45(5):443–461, August 2004.

[15] Kok Fong See, Ying Chu Ng, and Ming-Miin Yu. An alternative assessment approach to national higher education system evaluation. *Evaluation and Program Planning*, 94:102124, October 2022.

[16] Septi Andryana, Teddy Mantoro, Aris Gunaryati, and Alfarizky Esah Raffliansyah. Improving MCDM University Rankings through Statistical Validation Using Spearmans Correlation and THE Benchmark. *Journal of Applied Data Sciences*, 6(3), 2025.

[17] U.S. Department of Education, College Scorecard. Most Recent Institution-Level Data, November 2025.

[18] U.S. Department of Education, Federal Student Aid. School Data.

[19] U.S. Department of Education, National Center for Education Statistics (NCES). Integrated Postsecondary Education Data System (IPEDS): Fall, Winter, Spring Collection Cycles, 2021, 2022, 2023, 2024.

[20] U.S. Department of Education, Voluntary Support of Education. Voluntary Support of Education Data.

[21] U.S. National Science Foundation. Doctoral Degree Data, 2022.

[22] J. Fredericks Volkwein and Kyle V. Sweitzer. Institutional Prestige and Reputation among Research Universities and Liberal Arts Colleges*. *Research in Higher Education*, 47(2):129–148, March 2006.

[23] Xiodan Hu. Enrollment Challenges at Liberal Arts Colleges: A Literature Review. *Higher Education in Review*, pages 1–12, 2017.

[24] Murat Perit Çakr, Cengiz Acartürk, Ouzhan Alaehir, and Canan Çilingir. A comparative analysis of global and national university ranking systems. *Scientometrics*, 103(3):813–848, June 2015.

A. Codebooks

Here we present a more technical codebook. This is a record of the variable names we assigned to each variable as well as the description. We have also included table 1, showing the groupings we created for variables, again for the convenience of the reader.

Grouping	**Variables**
Applicant Demographics	Total Applicants; Total women applicants; Total men applicants; Percent of admitted students; Percent of admitted women; Percent of admitted men; Total Admissions yield; Admissions yield for women; Admissions yield for men
Applicant Test Scores	SAT math 25th, 50th, and 75th percentiles; SAT reading and writing 25th, 50th, and 75th percentiles; ACT composite 25th, 50th, and 75th percentiles
Alumni Donations	Average alumni donors; Average alumni solicited; Average total alumni donation; Average enrollment; Average percent alumni donors; average donation per alum; Average donation per student; Average gift 2022; Average gift per student
Enrolled Student Demographics	Enrollment; Percent of students American or Alaskan native; Percent of students Asian, native Hawaiian, or pacific islander; Percent of students African American; Percent of students women; Historically black college or university
Graduation Outcomes	Total graduation rate; Retention rate; Graduation rate 6y women; Graduation rate 6y men; Graduation rate 6y pell grant; Graduation rate 6y no loans; Percent doctorates; median earn working not enrolled graudates
Staff and Faculty	Total salaries and wages; Total staff; Instruction research FTE; library services FTE; education services FTE; management FTE; computer, engineering, and science FTE; service, legal, arts, and media FTE; healthcare FTE
Institution Finances	Instruction expenses; Endowment assets per student; Student service expenses; Endowment assets; Equity Ratio; Library expenditures per FTE; Salaries as percent of core expenses
Cost for Students	Total out of state price; Total in state price; Percent of students with financial aid; Average price for students on grant

Table 1. Copy of Table 1. Groupings for all variables. Semicolons separate variables, but for test scores, multiple are listed at once (SAT 25th, 50th, and 75th percentile represents three variables).

Variable	Description
ACT_composite_*	Score data; 25th, 50th, and 75th quantiles of ACT composite scores; 3YM
admissions_yield_*	Percentage data; Percent of students who were accepted to a school and attended; data for women, men, and combined; 3YM
applicants_*	Counts data; Number of students who applied to a school, percentage data available for women, men, and combined; 3YM
average_price_for_-students_on_grant	Dollars; Mean net price for students awarded any grants or scholarship aid; 3YM
avg_alumni_donors	Counts data; Mean number of alumni that donated after being solicited to donate; 3YM
avg_alumni_solicited	Counts data: Mean number of total alumni soliciated to donate; 3YM
avg_enrollment	Counts data; Mean number of students enrolled at an institution, enrollment from VSE data; 3YM
avg_gift_per_student_2022	Dollars; Mean nominal deferred gift for the year 2022 per student; 3YM
avg_giving_per_alum	Dollars; Mean donation per alumni solicited; 3YM
avg_giving_per_student	Dollars; Yearly donations divided by number of students; 3YM
avg_nominal_gifts_2022	Dollars; Mean nominal deferred gift total for the year 2022; 3YM
avg_percent_alumni_donors	Percentage data; Percent of alumni that donate after being solicited to donate; 3YM
avg_total_alumni_giving	Dollars; Mean total donation amount from the alumni that donated; 3YM
computer_engineering_-science_FTE	FTE staff; Full time equivalent Computer, Engineering, and Science staff; 3YM
education_services_FTE	FTE staff; Student and academic affairs as well as other education services full time equivalent staff; 3YM
endowment_assets_-per_student	Dollars per students; Endowment assets divided by year end FTE enrollment; 3YM
enrollment	Counts; Number of students enrolled for credit at an institution; 3YM
equity_ratio	Unitless; Ratio of net assets to total assets; 3YM
graduation_rate_6y_*	Percentage data; Graduation rate within 6 years. Contains data for all students as well as subpopulations for women, men, students without federal loans, and students with a pell grant; 2YM
healthcare_FTE	Counts; Full time equivalent Healthcare staff; 3YM
historically_black_-college_university	Unitless; Logical variable of if a institution is recognized as a historically Black college or university

instruction_expenses	Percentage data; mean percentage of instruction expenses as a percentage of core expenses; 3YM
instructional_research_-FTE	FTE staff; Full time equivalent staff functioning as librarians, curators, other teaching, and instructional support staff; 3YM
library_expenditures_-per_FTE	Dollars per student; Total library expenditures divided by 12 month FTE student enrollment; 2YM
library_services_FTE	Counts; Full time equivalent staff for librarians, curators, and other teaching instruction support staff; 3YM
management_FTE	Counts; Full time equivalent management staff; 3YM
MD_EARN_WNE_P10	Dollars; Median earnings of alumni that are working and not enrolled in school 10 years after undergraduate; 10YMedian
perc_doc20132022	Percentage data; mean percent of alumni that earned a Doctoral Degree from 2013-2022; 10YM
percent_admitted_men	Percentage data; mean percentage of male applicants that are admitted to a institution; 3YM
percent_admitted_total	Percentage data; mean percentage of applicants that are admitted to a institution; 3YM
percent_admitted_women	Percentage data; mean percentage of female applicants that are admitted to a institution; 3YM
percent_of_students_-financial_aid	Percentage data; mean percentage of students that recieve any kind of financial aid; 3YM
percent_students_-african_american	Percentage data; mean percentage of students that are African American; 3YM
percent_students_-american_alaska_native	Percentage data; mean percentage of students that are American Alaska Native; 3YM
percent_students_-asian_native	Percentage data; mean percentage of students that are Asian, Native Hawaiian, or Pacific Islander; 3YM
percent_students_women	Percentage data; mean percentage of students that are women; 3YM
retention_rate	Percentage data; mean full-time retention rate; 3YM
salaries_as_percent_of_-core_expenses	Percentage data; Mean percentage of salaries and wages as a percentage of core expenses; 3YM
SAT_math_*	Score data; 25th, 50th, and 75th quantiles of SAT math scores; 3YM
SAT_read_write_*	Score data; 25th, 50th, and 75th quantiles of SAT reading and writing scores; 3YM
service_legal_arts_-media_FTE	Counts; Total number of Community Service, Legal, Arts, and Media staff; 3YM
student_service_expenses	Percentage; Percent of funds for all core expenses allocated to student services; 3YM
total_in_state_price	Dollars; mean total tuition price for in-state students living on-campus; 3YM
total_out_of_state_price	Dollars; mean total tuition price for out-of-state students living on-campus; 3YM

`total_salaries_and_wages`	Percentage data; mean percentage of salaries and wages from total expenses; 3YM
`total_staff`	Counts; mean total number of full-time staff at a institution

TABLE 6. Quantitative codebook. These are the code names of every variable we used as well as more technical descriptions. Variable descriptions are of the form [UNIT]; [DESCRIPTION]; [TIME FRAME].

B. PCA Biplots

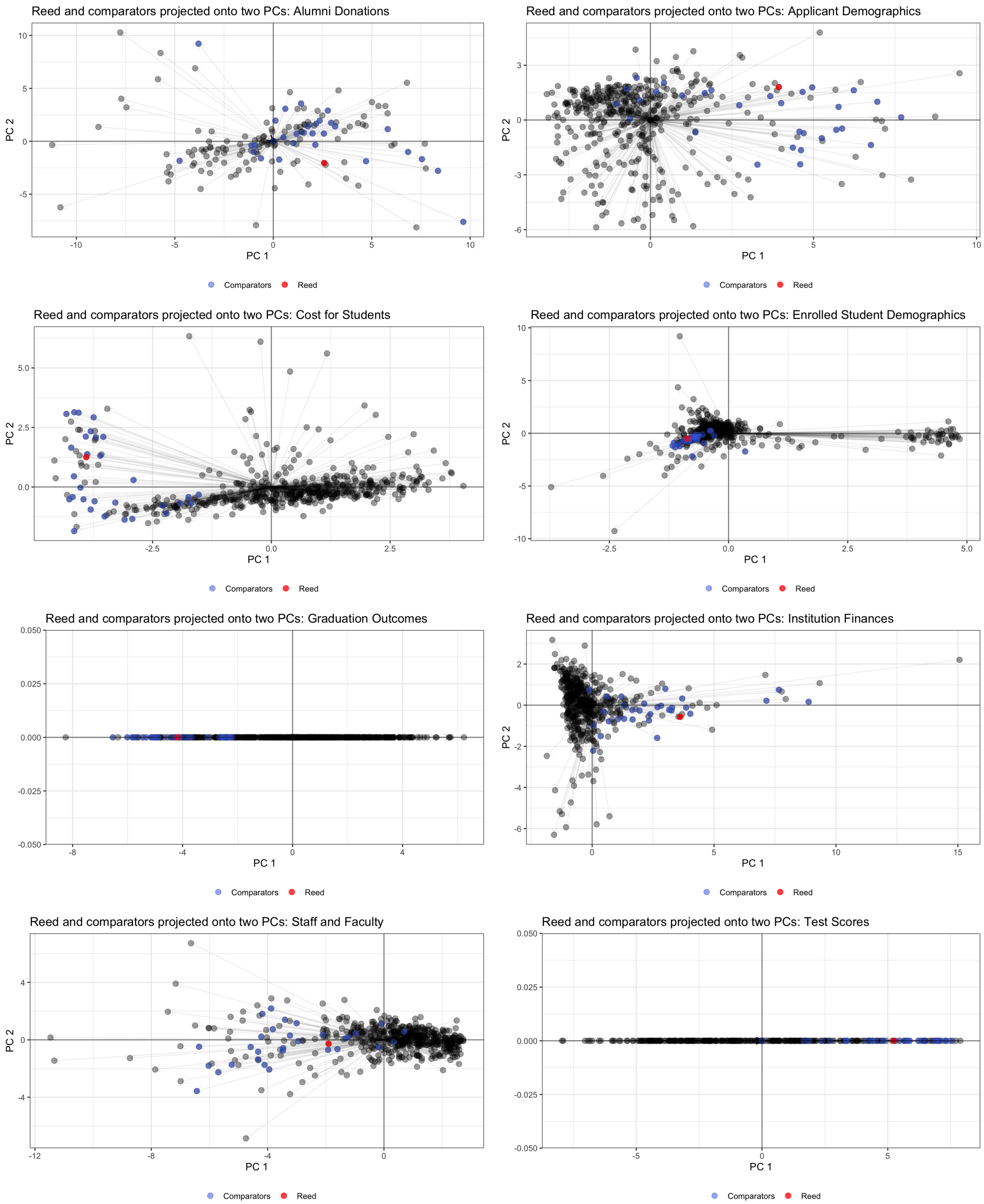

FIGURE 4. PCA one or two dimensional biplots for all groups. When groups had three principal components, we only showed two.

C. All Group Distances

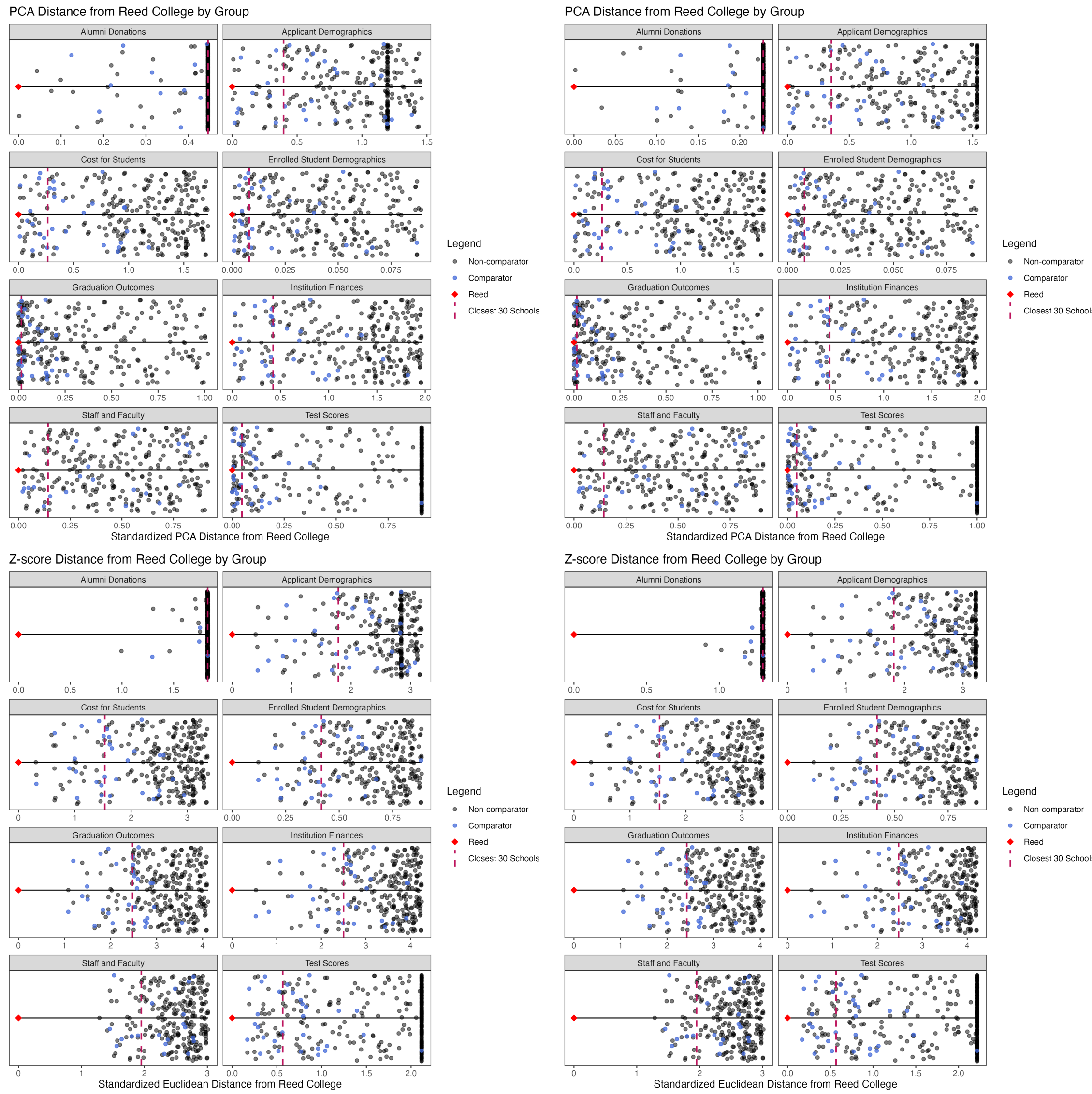

FIGURE 5. Group distance plots for for each of PCA mean, Z-score mean, PCA median, and Z-score median data imputation. The red diamond represents Reed College, the blue dots represents schools that are in the current comparator list, and black dots represent schools that are not in the current comparator list. The pink line represents the distance of the 30 closest schools to Reed College. NA values were imputed with the mean (left graphs) and median (right graphs) of a variable.

D. All Distance Matrix Plots

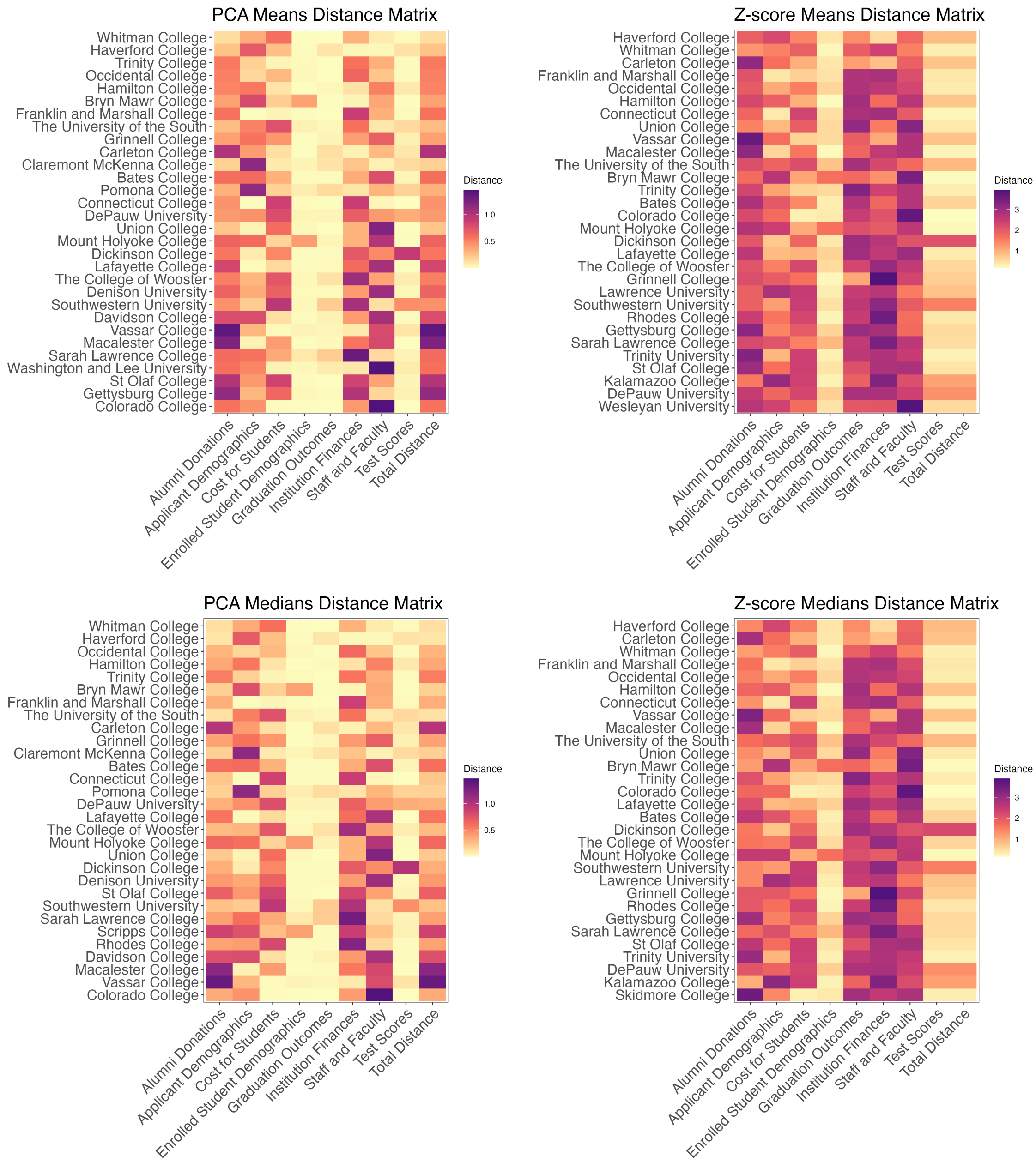

FIGURE 6. PCA and Z-score with means data imputation (top) and medians data imputation (bottom) showing the distances in each coordinate for the top thirty schools. Scales are automatically determined to incorporate all values. The x-axes are the different groups, and the y-axes are the top thirty schools.

E. Shiny App Specification

There are two versions of the shiny app. There is the public shiny app referenced in [13]. This has the app as well as a number of csv files representing already completed analyses. These have only a few summary statistics and none of the original data. These are intended to be shared widely, as the VSE data is not public. Alongside these is a shiny app which can easily be run.

Along with this, there is a private tarball (`.tar.gz` file) containing all of the original data alongside all of the scripts. This will be distributed to Michael Pearce and Mike Tamada. From these, one can repeat our analyses and regenerate our data. Equally, it is possible to modify the analyses and as long as the structure of the data doesn't change, it should compile the data. The app is here as well and if the data has been compiled, the app should run easily.

To make an analysis available to the app, a csv file with some values for each group for every school and zero `NA` values should be added to its folder. Currently, there are the eight groups discussed earlier hard-coded into the analysis and the app. These can be changed provided the app script is expanded to accommodate new groups.

As of right now, the app is not online. It may be run from the RStudio application or anything else that can natively run shiny apps. Alternatively, it may be run by calling `make run` in a terminal prompt with a working directory in the folder with the app. We will plan on hosting it on a server run by Reed, allowing only people with a Reed connection to connect to the app and not requiring them to interact with any of the back end.