

# Reference Group Formation Using the Nearest Neighbor Method

by Gerald W. McLaughlin, Josetta S. McLaughlin, and Richard D. Howard

## INTRODUCTION

In a dynamic political environment, the analysis applied to as sensitive an issue as peer comparisons must necessarily reflect the adjustments and compromises that are part of the political process. In return, decision making that draws from sound analysis is more likely to avoid the manipulations of the purely political process. (Weeks, Puckett, and Daron 2000, p. 20)

American higher education has long been regarded as a major economic and social engine of change. To maintain this role, institutions must themselves make appropriate changes that keep them aligned with the needs of a dynamic environment. Accomplishing these changes requires institutions to create a roadmap that informs their administration about how they are performing internally and externally relative to goals and to other institutions that are operating in the higher education sector. Developing reliable benchmarks is generally the operational choice for this activity. Benchmarking can also help institutions create an ongoing sense of urgency by alerting them to a need for change (Qayoumi 2012). However, to set appropriate goals and objectives against which to benchmark, an institution must first develop a set of reference institutions.

Various types of reference groups are commonly used by higher education institutions to benchmark performance outcomes (McLaughlin and McLaughlin 2007). In fact, interest in forming groups for this purpose is not new; the exploration of statistical methodologies that can be used to group institutions began more than 30 years ago (Brinkman 1987; Brinkman and Teeter 1987; Korb 1982; McLaughlin and McLaughlin 2007; Teeter and Brinkman 1987; Terenzini et al. 1980). The primary objective was then, as it is now, to find an appropriate method for comparing the performance of one institution against norms developed using a relevant group of comparator institutions. Reference group formation thus evolved as an alternative to the “industry norm” system used by many business sectors to evaluate financial performance. Over time, creating norms through reference group formation has in fact proven to be a viable tool for benchmarking an institution’s performance outcomes.

*Over time, creating norms through reference group formation has in fact proven to be a viable tool for benchmarking an institution’s performance outcomes.*

Reference group formation is a response both to internal planning needs and to external pressures for accountability in higher education (Bender and Schuh 2002). As noted by Trainer (2008, p. 17):

In this age of accountability, transparency, and accreditation, colleges and universities increasingly conduct comparative analyses and engage in benchmarking activities. Meant to inform institutional planning and decision making, comparative analyses and benchmarking are employed to let stakeholders know how an institution stacks up against its peers and, more likely, a set of aspirant institutions—those that organizational leaders seek to emulate.

The importance of reference groups has continued to intensify over recent decades due to changing perspectives about how higher education should fulfill its mission. There are increasingly calls for higher education institutions to report performance outcomes to such stakeholders as the government (especially for those institutions that receive state and federal funds), accrediting agencies, bond agencies, and other financial organizations evaluating institutional financial stability (Gaither, Nedwek, and Neal 1994; Townsley 2002). Initially, groups formed around predetermined criteria such as athletic conferences, state institutions, or faith-based institutions (Teeter and Brinkman 1987). Over the last few decades, the emphasis has shifted away from predetermined groups for all purposes to other types of groups. For example, “aspirational” groups that enable administrators to benchmark performance against institutions that have attributes that are desired (but not yet realized) by the focal institution are increasingly used to set a strategic and future direction (McLaughlin and McLaughlin 2007).

Several methods for forming homogenous groups have gradually found acceptance both inside and outside of higher education (Blankmeyer et al. 2010; Kerschbaum 2008) with the statistical tools of choice including some form of cluster analysis (Reiss et al. 2010). Nearest neighbor methods, which are closely related, have also gained acceptance (Borden 2005; Teeter and Christal 1987). Though the outcomes are sometimes difficult to explain to stakeholders, both cluster analysis and nearest neighbor methods have been proven to be useful. Potential problems have been identified (e.g., comparability, substitutability, the additive attributes of some procedures), but the relative objectivity of these statistical methods has provided a sound foundation for creating norms against which to benchmark (McLaughlin and McLaughlin 2007; Secor 2002; Xu 2008).

## CHOOSING METHODS FOR REFERENCE GROUP FORMATION

Reference group formation requires development of a systematic and transparent process that ultimately answers questions concerning which variables are appropriate for selecting reference institutions and what analytical tool is appropriate for use with these variables. Finding the answers can be accomplished using the seven steps identified below as a framework:

1. Clarify the purpose
2. Determine the composition
3. Select a methodology
4. Identify measures
5. Collect and analyze data
6. Determine acceptable magnitude of difference
7. Present results and adjust the process

*Reference group formation requires development of a systematic and transparent process.*

Though these steps suggest linearity, they do not always represent a linear sequence. The completion of one step frequently results in an iterative, but hopefully heuristic, cycle of revisiting previous steps while simultaneously moving to the next step in the sequence. Nevertheless, each step is discussed below in sequence.

### STEP 1: CLARIFYING THE PURPOSE FOR DEVELOPING AND USING THE REFERENCE GROUP(S)

Any initiative that requires formation of reference groups for benchmarking can have significant political, social, and economic impacts. These impacts potentially affect (positively or negatively) the professional status of faculty, administrators, staff, and, ultimately, students. As such, it must be recognized up front that reference group formation may be influenced by political agendas from across and often beyond the campus. In recognition of this influence, the group formation process must start with a clear statement of the purpose and the intended use of institutional comparisons.

The purpose for reference group formation can be extremely broad (e.g., comparing overall institutional effectiveness with other “peer” institutions) or extremely focused (e.g., setting goals for faculty research funding). Comparisons can focus on general categories such as salaries, staffing, adequacy of funding, and expenditures or on assessing outcomes such as graduation, debt, and student engagement (Pike and Kuh 2005). Beyond the traditional foci, political considerations may require creating benchmarks for specific areas of concern (e.g., diversity, sustainable development) (McLaughlin and McLaughlin 2007).

### STEP 2: DETERMINING THE TYPE, SIZE, AND NUMBER OF REFERENCE GROUPS TO FORM

The purpose of the benchmarking initiative(s) determines whether to create one general set of “peers” or multiple sets of “comparator” institutions (e.g., one for salary comparisons, one for retention and graduation comparisons, one for financial comparisons). In general, the larger and more complex an institution, the more likely it is that multiple comparison groups will be necessary. The smaller liberal arts college may need only a single group. The purpose will also impact decisions regarding the size of the groups. These issues are discussed below.

- **TYPES OF REFERENCE GROUPS.** Once the purpose of the benchmarking initiative has been determined, different types of reference groups can be created (McLaughlin and McLaughlin 2007). The most commonly formed group is the *peer group*. Institutions in this group are *similar* on most primary or key attributes (e.g., size, mission, student characteristics, curricula, fiscal and personnel resources).

A second reference group type is the *aspirational group*. Institutions in this group, sometimes referred to as “preferred peers,” possess one or more attribute(s) that the focal institution desires to possess but has not yet attained (e.g., higher status, greater resources, higher graduation rates). The institutions have similar characteristics on other measures. Institutions frequently identify this type of group using popular college and university rankings such as *U.S. News & World Report’s* “Best Colleges.”

A third reference group type is the *competitor group*. Institutions in this group compete with the focal institution for specific resources. For example, a competitor of interest would be the institution in which a student enrolls when (s)he is accepted to but does not enroll in the focal institution. Several organizations

collect data on where “students go” after acceptance (e.g., the National Student Clearinghouse, ACT). One factor complicating data collection in this example is that competitors for students and other resources may not be higher education institutions. The primary competitor for students could be military service, a local business, a large corporation, or an industry training program.

A fourth reference group type is the predetermined group. As noted earlier, these groups already exist for other purposes (e.g., faith-based institutions, athletic affiliation, jurisdictional affiliation). Classification groups—such as those formed by the Carnegie Foundation classification process—also fall under this category. The Carnegie Classification system is used extensively in rankings and national studies (e.g., College Results Online, *U.S. News & World Report* rankings, the NACUBO Benchmarking Tool, American Association of University Professors [AAUP] salary studies). (For a more extensive discussion of types of reference groups, see McLaughlin and McLaughlin 2007; Teeter and Brinkman 1987, 1992.)

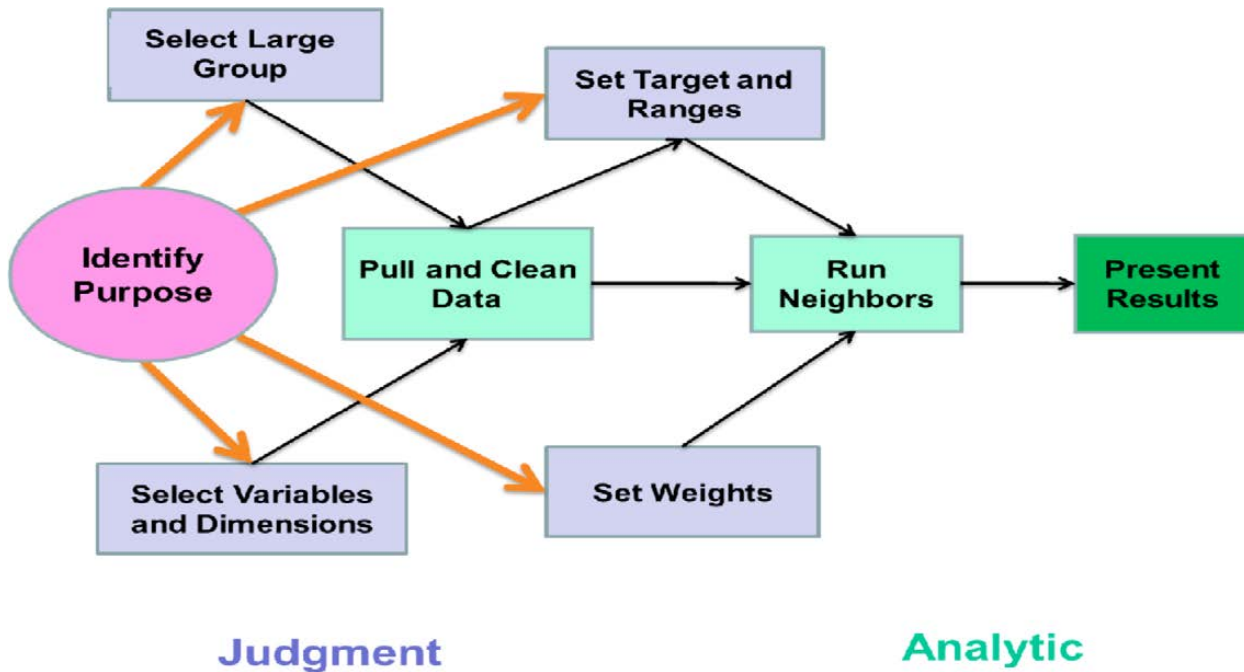
- **SIZE OF REFERENCE GROUP.** The appropriate size of a reference group is determined by its intended use and whether one or multiple groups are being identified. For example, a larger group of regional institutions—25 or 35—may be used to create norms for multiple purposes (e.g., to both benchmark and develop goals and objectives) or for different types of performance outcomes (e.g., a retention and graduation rate benchmarked against the reference group’s median or an aspirational goal benchmarked against the reference group’s 75th percentile). In some cases, institutions may want to form a group of four or five very similar institutions as peers and a second group of four or five institutions that represent an aspirational group. Unfortunately, the smaller the group(s), the more likely it is that there will be political objections concerning the group’s appropriateness.
- **NUMBER OF REFERENCE GROUPS.** If the purpose of constructing reference groups is broad general use, then a single reference group will likely be sufficient for benchmarking purposes. Also, constructing larger groups against which to benchmark makes it more likely that data will be available for a sufficient number of the preferred comparison institutions. Data availability for some measures is a problem in that many of the data sources are membership organizations (e.g., CUPA-HR, the Consortium for Student Retention Data Exchange [CSRDE], the National Survey of Student Engagement [NSSE]) and their members may be only a subset of the institutions of interest.

Many other factors (e.g., human resources, student characteristics, financial viability) can affect the decision of whether to use different groups for different benchmarking initiatives. To focus efforts and control for availability of resources, institutions within a similar sector (e.g., public or private not-for-profit; private for-profit; urban/rural) may be used. However, when looking at student characteristics, other institutions with similar curriculum profiles and a similar balance of attributes (e.g., residential/commuter, graduate/undergraduate/professional, minority/ethnic student characteristics, socioeconomic status) can be more useful. If the intended purpose is to benchmark against competitors for students, then the reference group likely will contain a larger number of institutions that are in close geographic proximity to the focal institution.

**STEP 3: SELECTING A METHODOLOGY FOR FORMING THE REFERENCE GROUP(S).**

The methodology for reference group formation generally follows the pattern shown in figure 1. The primary methodologies for forming reference group(s), as noted earlier, are judgment, analytics, or use of preexisting groups based on classification (e.g., athletic conference). Judgment typically builds on the expert opinion of an institution’s key stakeholders. As noted, this methodology tends to be fairly simple but politically sensitive and quite contentious when there are competing interests. On the other hand, analytics provide a more objective alternative, but results can be difficult to explain. Most reference group formation initiatives thus use some combination of judgment and analytics.

Figure 1 **Methodology for Forming Reference Groups**



Analytics using cluster analysis define a large group of institutions within a multidimensional space formed from selected variables (i.e., measures typically related to the purpose for which the group is to be used). The measures are traditionally converted to standardized measures after which a composite metric is formed for the distance between each pair of institutions (Terenzini et al. 1980). Decisions are made concerning how to standardize variables, how to ensure consistency in measures (e.g., does a dollar in salary count more than a dollar in tuition), whether variables should reflect magnitude or relative magnitude, and whether variables should be based on size (e.g., number of faculty, number of students) or on ratios (e.g., students per faculty, average salary per faculty). When using cluster analysis, no definitive rules exist for the number of clusters that is appropriate. Though more objective than some other methods, some judgment is required when the institution of interest is on the outer boundary of a cluster, making it potentially more similar to institutions in another cluster. (For additional information, see Nisbet, Elder, and Miner 2009; StatSoft.com n.d.)

The nearest neighbor approach, in contrast to cluster analysis, defines the institution of interest as the centroid in the space defined by the variables. Other institutions are evaluated on their distance from the target institution. This methodology has several variations; however, the distance can be measured with metrics that are typically standardized

and then weighted (Weeks, Puckett, and Daron 2000). In some instances, a specific set of characteristics is selected, and those institutions without these characteristics are excluded. The advantage to using nearest neighbor methods is that the institution of interest is at the center of the most similar institutions available given the variables selected for the analysis. The disadvantage in using this method is that there is no clear number of institutions that need to be included in the analysis. Determining an appropriate number of institutions thus requires a form of judgment and a continuing discussion about the purpose of forming the reference group.

#### STEP 4: IDENTIFYING MEASURES FOR FORMING REFERENCE GROUPS

Measures are those institutional attributes that support identification of comparable reference institutions for the purpose of benchmarking. The measures used to implement the group formation methodology must possess the appropriate attributes for addressing the purpose for which the group is to be used (Wang and Strong 1996). First, the chosen measures must add value to the process of group formation by defining the measurement domains of interest (e.g., benchmarking student retention rates requires measures of student characteristics; management effectiveness requires financial comparability measures; faculty productivity requires measures of instructional workload and scholarship activity). If predetermined groups are used, measures may be as simple as identifying the geographical region or athletic conference. If analytics or judgment (or a combination of the two) is used, the choice of measures is more sophisticated. Measures will be aggregates and surrogates for the underlying characteristics or constructs being measured. This requires that two criteria be met. First, measures must have a direct connection to the characteristic of interest, reflect the goals or objectives of the strategic process, and be fairly robust against inappropriate manipulation. Second, measures must possess several basic characteristics—reliability, validity, timeliness, sufficiency, relevancy, and transparency. They should also be economically available (Health Information and Quality Authority 2011; McLaughlin, Pavelka, and McLaughlin 2005; Weeks, Puckett, and Daron 2000). Each characteristic is described below.

- **RELIABILITY.** Reliability is an essential characteristic of measures. To be reliable, measures must be objective, stable, and internally consistent (Howard, McLaughlin, and McLaughlin 1989). If the measure is objective, two individuals collecting data on the same measure should get the same number. If the measure is stable, the definition of the measure should be the same over time. If the measure has internal consistency, the measure should represent the core component of interest. For example, a measure must be defined in the same way across all institutions in the reference group. Financial data are particularly problematic since the three major sectors of higher education (i.e., public, private not-for-profit, and private for-profit) are required to use different accounting and reporting standards. These requirements are compounded by a range of other issues that must be addressed (e.g., categorization of specific costs such as instructional technology, classification of various types of expenditures such as financial aid).
- **VALIDITY.** A measure is valid only if it accurately reflects what it is supposed to reflect. Measures that are not reliable cannot possess validity. Measures must be correctly defined, coded, understood, and interpreted. Beyond definition, measures must be appropriate for generalizing interpretations to relevant situations both in the present and possibly in the future. The most significant threat to validity is bias (i.e., a prejudice in favor of or against a particular interest). The danger inherent in bias is that the measure will systematically misrepresent what it claims to measure. For example, the traditional graduation rate does not include transfer students or those who graduate after 200 percent to degree time, thus misrepresenting the actual graduation rate for certain types of colleges and universities. (Note that it would not be appropriate to

compare institutions with different definitions for graduation rates, e.g., four-year vs. two-year institutions.) Similarly, the ratio of degrees divided by enrollment gives a number that is too high for institutions with shrinking enrollments or a larger proportion of transfers. To address problems of bias, multiple measures can be used to measure an important characteristic (e.g., socioeconomic status [SES] as measured by combining income, occupation, and education level).

- **TIMELINESS.** To be useful, measures must be available when needed to support the reference group formation. The different external and internal measures describing institutions used in reference group formation should reflect the same time period. For example, dividing “2012 enrollment” by “2010 faculty” is not appropriate.
- **SUFFICIENCY.** The measures describing key aspects of the institution must be sufficient (i.e., must represent the major characteristics of the institution, such as size, highest degree, size and setting, selectivity, and financial capability). For example, if the purpose of constructing a reference group is to examine retention, a measure such as highest degree awarded may not be needed but measures of student financial status and academic ability are needed. The level of detail in the measure is also important. Most measures from federal and state data sets are available at the institutional level, although some are aggregated at the system or state level. Limited measures are available at the discipline level.
- **RELEVANCY.** Measures should be relevant to the purpose for which the reference group is to be used. To be relevant, measures must increase the quality of understanding about the situation and about the nature of the decision being made. Measures that do not contribute to understanding should be eliminated. Based on the principle of parsimony and the belief that humans can perceive only a limited number of items, the use of excessive (i.e., non-contributing) measures should be avoided so that attention can be focused on important aspects of the question.
- **TRANSPARENCY.** Measures are transparent if they are defined and described in a manner that is both interpretable and understandable by stakeholders. Definitions should be publically available, logical, and consistent with “traditional wisdom.” For example, traditional wisdom suggests that selectivity and average SAT or ACT represent entering student ability. While other measures might be used, there are definite advantages to using traditional and transparent measures that are understood by the public. Definitions of transparency include what is referred to as content or face validity.
- **ECONOMICALLY AVAILABLE.** Measures are economically available for reference group formation if the benefit of using the measure exceeds its costs of collection, analysis, and use. The most economic measures are those currently and readily available. Costs are increased if data for the comparison institutions must be acquired through development of a data exchange. Fortunately, external data are available for most institutions to use in reference group formation when measures are needed for common goals such as growth (Trainer 2008).

As noted, federal, state, and private databases contain usable measures (e.g., federal IPEDS data, state college and university fact books, private-sector *U.S. News & World Report*). National databases are available from organizations such as GuideStar (financial data on private institutions) and the U.S. Department of Education College Navigator. Finally,

specialized membership databases like CSRDE (retention data), CUPA-HR (salary data), and the Delaware Study (instructional workload and costs by department) contain useful data. The major federal database, however, is the IPEDS Data Center, which is available along with multiple tools through the National Center for Education Statistics (NCES).

#### **STEP 5: COLLECTING AND ANALYZING THE DATA TO FORM REFERENCE GROUPS**

The primary federal entity collecting data on higher education institutions is the NCES. Its IPEDS data are clearly its main source of data, and its tools are publicly available online for the reference group formation initiatives that make performance benchmarking feasible. Several other organizations (e.g., NACUBO, CSRDE) provide tools for members. In addition, websites designed to help students select an institution provide “search” options that result in a set of institutions that mimic a reference group. Two examples are NCES’s College Navigator and the Carnegie Foundation classification system. While these resources offer a starting point for reference group formation, they may not provide all needed data. The data available through these sites are nevertheless economically available and appropriate for use with nearest neighbor methods.

#### **STEP 6: DETERMINING THE ACCEPTABLE MAGNITUDE OF DIFFERENCE WHEN FORMING REFERENCE GROUPS**

Discussions about reference group formation generally focus on the amount of homogeneity, or similarity, represented by the institutions. However, questions must also be answered concerning how different an institution needs to be from the focal institution on specific measures in order to be excluded from a group. For example, if a potential comparator institution has a hospital while the focal institution does not, a decision must be made concerning whether to include the comparator institution in the group. Similarly, decisions must be made in cases where one institution focuses on undergraduate instruction while the focal institution and all other institutions in the group have substantial doctoral programs.

Determining the importance and magnitude of a difference is also related to the importance and weighting of a measure (i.e., variable). Not all variables are equally important in the reference group formation process. A number of different approaches can be used to indicate the importance of a variable. Variables that are considered to be important can be weighted by entering them multiple times into a nearest neighbor methodology or can be standardized to increase their functional weight. For example, weighting used by Lang (2000) started with the basic categories of enrollment, financial, library, demographic context, and degrees awarded from which 23 individual aspects of the institutions were identified. Different sets of weights were created from these 23 measures—one set of weights represented a General Slate perspective; one set of weights a Research Slate; one set of weights a Compensation Slate; and one set of weights a Government Ability to Pay Slate. Full-time equivalent (FTE) enrollment was weighted at 5 percent in the General Slate and Government Ability to Pay Slate, 2 percent in the Research Slate, and 0 percent in the Compensation Slate. The percentages were then multiplied times the standardized differences between various focal institutions and the other institutions under consideration for inclusion.



## STEP 7: PRESENTING RESULTS AND ADJUSTING THE PROCESS WHEN FORMING REFERENCE GROUPS

Presenting the analytical results of the reference group formation process generally results in additional iterations based on stakeholder feedback. The iterative process typically brings judgment to bear on all steps of the process by combining analysis and judgment. In fact, if reference groups are being formed for an applied purpose, then it is highly unlikely that the results of the reference group formation process will be fully quantitative or accepted after the first iteration. This will be demonstrated in the following case example.

### APPLICATION: THE CASE STUDY

This case study demonstrates reference group formation using the nearest neighbor method. The focal institution for this study is a southeastern land-grant university with very high research activity and numerous doctoral programs. Institutions identified for inclusion in the group formation process confer bachelor's, master's, and doctorate degrees.

### PURPOSE

The purpose for reference group formation is to identify similar institutions that can be used for general benchmarking purposes. In this case, there is no specific focused agenda item that requires developing a specialized reference group. The intent is to use historically common institutional attributes that make a general reference group of value for benchmarking. In addition, the final reference group must support goal setting for multiple benchmarking activities.

### NUMBER AND SIZE

With respect to the number of reference groups to form and the size of each group, the methodology chosen is flexible and can be used to create either multiple reference groups or a single reference group to support benchmarking initiatives. The size of the reference group can range from very small to very large. Because the intent in this case is to create a group for multiple uses, a relatively large group of manageable size (in the neighborhood of 25 or 30) is created.

### METHODOLOGY

Nearest neighbor methods are used to create the reference group. As noted, the focal institution is classified as high research with numerous doctoral programs; institutions for possible inclusion must confer bachelor's, master's, and doctorate degrees. Institutions that do not offer doctorate degrees, institutions outside of the United States and the District of Columbia, and private for-profit institutions are excluded. It is highly unlikely that these institutions would be accepted by stakeholders as comparable to a major research university. Institutions are also required to be Title IV eligible. After exclusions, an initial group of 559 institutions is available. Eleven institutions are then removed because of excessive missing data.

**MEASURES**

Measures are identified through multiple iterations based on key operational areas in the focal institution and on data availability. The categories of interest and items chosen for analysis are shown in figure 2.

Figure 2 **Categories and Items Used to Identify Nearest Neighbors**

<b>Institutional Characteristics</b>	<b>UG Market Characteristics</b>	<b>Student Characteristics</b>
Population Density	<i>*FTE Students</i>	% White Students
Region	UG Freshmen Applicants/UG HC	% UG as Female
Carnegie Basic	UG (IS) Tuition and Fees	Dorm Capacity as % FT UG
Carnegie UG Profile	% Discount Rate (Fees)	% UG as Full Time
Carnegie Enrollment Profile	% FT-FT DS Accepted	% UG Entering in FT-FT DS Cohort
<i>*Carnegie Size and Setting</i>	Yield of FT-FT DS	% FT-FT DS Cohort with Pell Grants
Control	Freshman Retention Rates	Student Services \$/FTE Student
Hospital	6-Yr Graduation Rates	% UG 25 Years and Older

<b>Academic Characteristics</b>	<b>Curriculum Characteristics</b>	<b>Financial Characteristics</b>
IPEDS Student/Faculty Ratio	First Prof and PhD's as % Degrees	Net Tuition + State Dependency /Core Revenues
% FTE Staff as Faculty	<i>*Engineering as % Bachelors</i>	Tuition and Fee and State Revenue /FTE Student
<i>*Research &amp; Service \$/FTE Faculty</i>	Educ/Leisure/Family Science as % Bachelors	Endowment \$/FTE Student
% Full-Time Faculty as White	<i>*Other STEM as % Bachelors</i>	Net Income Ratio
% Full-Time Faculty as Female	Bus/Pub Admin/Legal/ Communications as % Bachelors	Financial Viability
Average Faculty Salary	Applied PhD's as % (First Prof + Doctoral)	Primary Reserve Ratio
% FTE Faculty as Tenure Track	<i>*Educ/Leisure/Family Science as % Graduate</i>	Return on Net Assets
Instruction and Academic Support \$ /FTE Student	Technology and Health Science as % Degrees	% Change in Endowment

Abbreviations: DS: degree-seeking; FT-FT: first-time, full-time; FTE: full-time equivalent; HC: head count; IS: in-state; UG: undergraduate.

Note: Italicized text and "\*" indicate variables that are weighted 2 (i.e., more important).

Variables are assigned different weights by multiplying those variables believed to be more important by 2. In reference group formation, certain variables (such as retention rates) can on occasion be used as input variables even though the intent may be to use the reference group for benchmarking against an output measure.

## COLLECTION AND ANALYSES OF DATA

Data for analyses are obtained from the online IPEDS Data Center (<http://nces.ed.gov/ipeds/datacenter/>). The appropriate institutional .uid and variable .mvl files are developed and used to extract the data. Financial data, programmatic data based on degrees conferred, and general institutional and staffing data are extracted as three different data sets and converted to Excel spreadsheets. After the data are sorted by the Institutional Federal ID number (UNITID), they are copy-pasted into worksheets of a master Excel workbook. Data from the three worksheets are then combined on a fourth worksheet into a balanced scorecard using formulas that bring the data from the individual worksheets into the scorecard. Once the scorecard is constructed, each institution's distance from the focal institution is computed for each measure. These distances are standardized, weighted, and then summed to form a *Proximity Index* for each institution. These steps are described in more detail in the following discussion.

## MAGNITUDE OF DIFFERENCES AND ANALYSIS

The analysis requires calculating differences and adjusting for the magnitude of differences in computing the proximity of each institution to the focal institution. The following procedures are used to create the Proximity Index:

1. For each institution, the measures listed in figure 2 are compared to the focal institution and assigned a difference score of "0," "1," or "2." Zero indicates that the institution is the "same" as the target institution on the item; 1 indicates that the institution is "similar" to the target institution on the item; and 2 indicates that the institution is "different" from the target institution on the item.
2. For categorical items (e.g., institution type), judgment is used to determine the degree to which an institution is the "same," "similar," or "different." Categorical items include institutional type variables. For example, in the case of this major research land-grant university, designation under the basic Carnegie category of "Very High Research/Doctoral" is considered to be the same, "High Research/Doctoral" is considered to be similar, and all other institutional categories are considered to be different.
3. For continuous items (e.g., expenditures), basic differences are established using the standard deviation of the item. The following definitions are used:

Let  $\Delta = |\text{Target Institution minus Other Institution}|$ , then

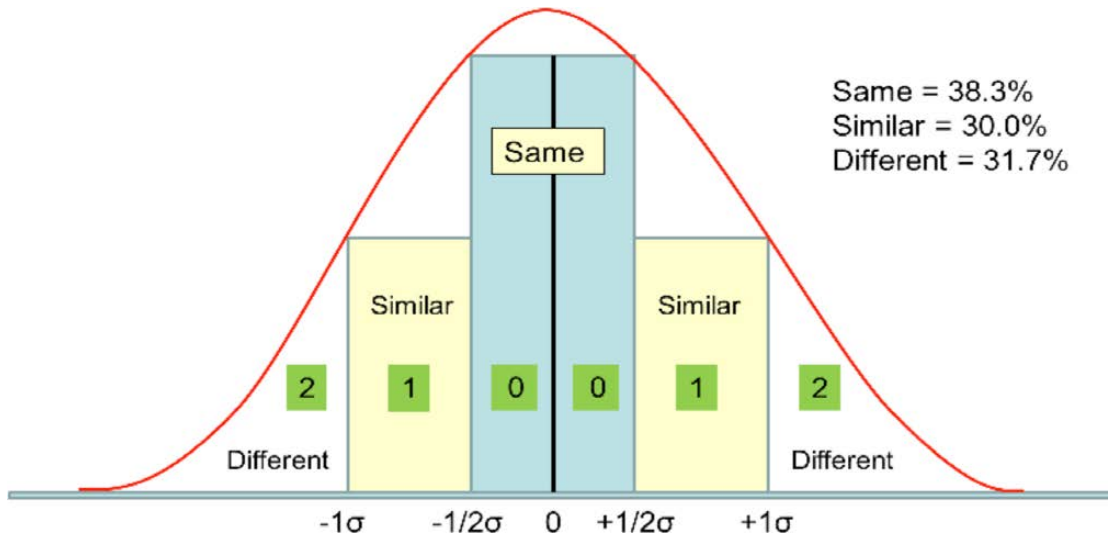
Same = If  $\Delta \leq \frac{1}{2}$  Standard Deviation (SD) then  $X_i = 0$ ;

Similar = If  $\frac{1}{2}$  SD  $< \Delta < 1$  SD then  $X_i = 1$ ;

Different = If  $\Delta \geq 1$  Standard Deviation then  $X_i = 2$ .

Using Excel, a gap score is calculated that indicates how similar to or different from the focal institution another institution is. In cases where a variable is highly skewed, an adjustment is made in the width of the gap to result in an approximately equal distribution of institutions across the categories. The result is a better and more balanced proportion of observations designated as "same," "similar," and "different." As shown in figure 3, if the scores on an item are normally distributed, then approximately 1/3 of observations will fall into each category.

Figure 3 Preferred Distribution of Difference Scores



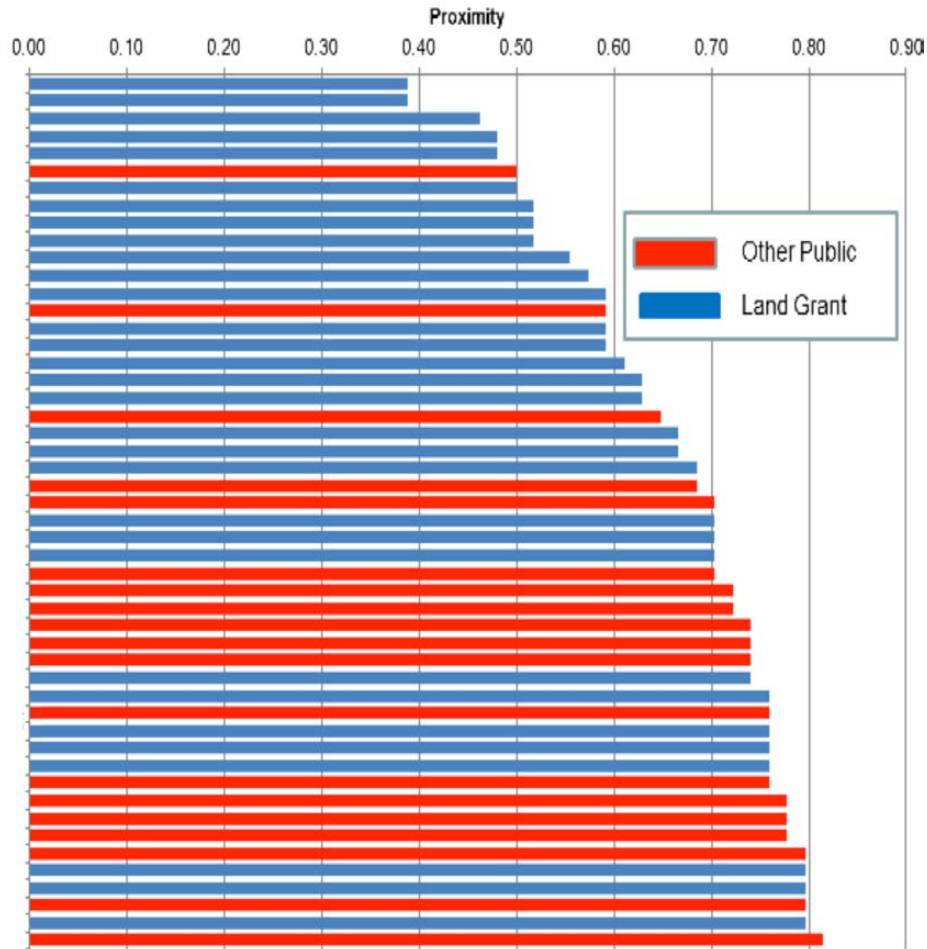
4. For each institution, the similarity score is weighted and summed across the 48 items shown in figure 2. This sum of the 48 scores is then divided by the sum of the weights so that each institution gets an aggregate score between zero and two (referred to here as a Proximity Index).

## RESULTS AND ADJUSTMENTS

Results from using the nearest neighbor methods are presented to stakeholders, and the process is adjusted based on user feedback. Institutions are designated as “similar” based on their overall proximity to the focal institution and on their proximity in terms of the 48 specific measures used to compute the overall proximity. Figures 4–7 show results that demonstrate the information created through the use of the methodology.

Figure 4 shows the proximity of the most similar 50 institutions to the focal institution. Institutions with a Proximity Index less than .5 are considered to be the “same,” institutions with a Proximity Index from .5 to 1.5 are considered to be “similar,” and institutions with a Proximity Index greater than 1.5 are considered to be “different.” The range of institutions considered for inclusion in the reference group have proximity scores between .39 and .82. The visualization of the results enables strategic planners to decide how many institutions might be appropriate for reference groups of different sizes. For example, it is interesting to note that six of seven institutions that have a Proximity Index of .5 or less are land-grant universities. The seventh, while not a land-grant university, is a major public research institution. These seven institutions could make up a small reference group for comparative purposes in benchmarking specific functional areas (e.g., administrative processes, student life programs).

Figure 4 Proximity of Closest 50 Institutions to the Focal Institution



Note: Proximity Index of < 0.5 = "same" as focal institution; 0.5–1.5 = "similar" to focal institution; > 1.5 = "different" from focal institution.

Figure 5 uses a spiderweb chart to show the proximity of types of institutions (e.g., land-grant universities, other public universities, private universities) to the focal institution on the specific characteristics of interest identified in figure 2. The type of institution and the average total Proximity Index for each type are included. The focal institution is at the center of the web. The blue line represents land grants; red, other public universities; and green, private universities. As expected, the greatest differences in proximity are in type of institution. With respect to curriculum, private universities are actually more similar to the focal institution than are other public universities.

Figure 5 Similarity of Three Types of Institutions to a Southeastern Land-Grant Research University on Six Measures

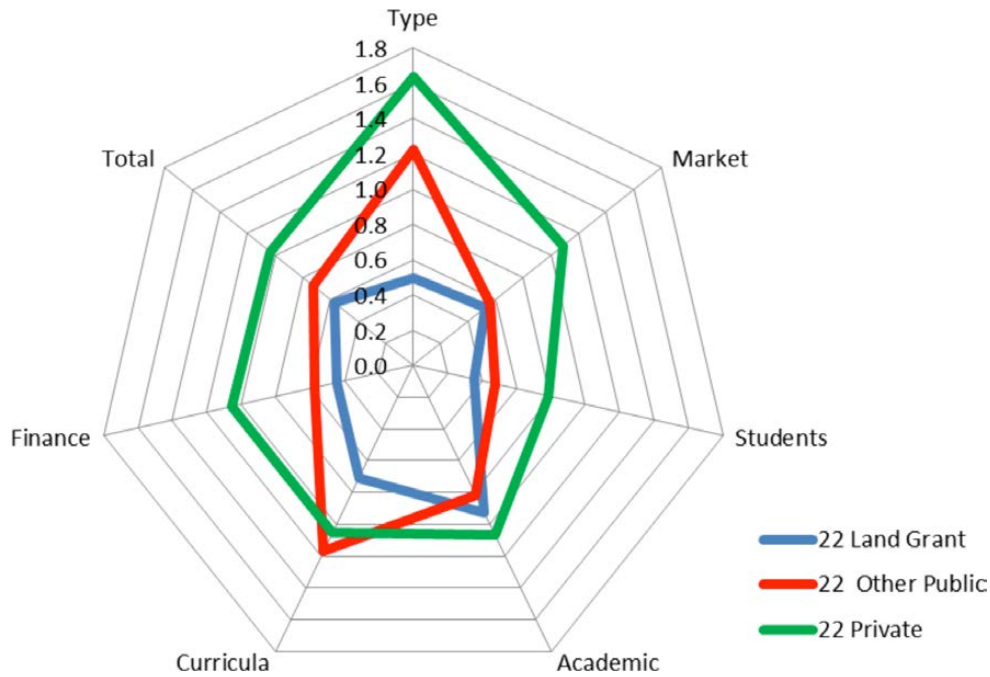


Figure 6 uses a spiderweb chart to show the similarity of institutions to the focal institution in a specific area of interest—curriculum characteristics. The analyst is able to drill down into the data to better understand in which areas of interest the variations occur. Results indicate that in terms of curriculum, land-grant, other public, and private universities tend to differ from the focal institution on two items—“engineering” and “other STEM” as a percentage of bachelor’s degrees awarded.

Figure 6 **Similarity of Three Types of Institutions to a Southeastern Land-Grant Research University on Eight Curricula Characteristics**

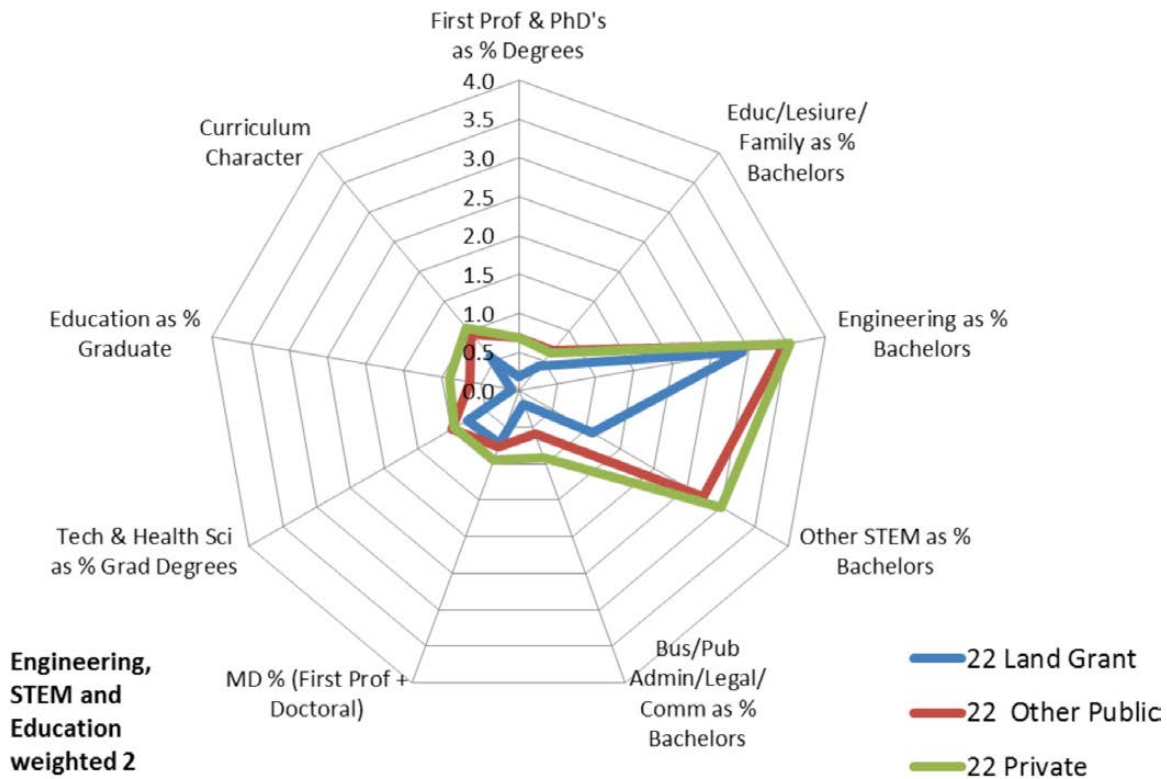
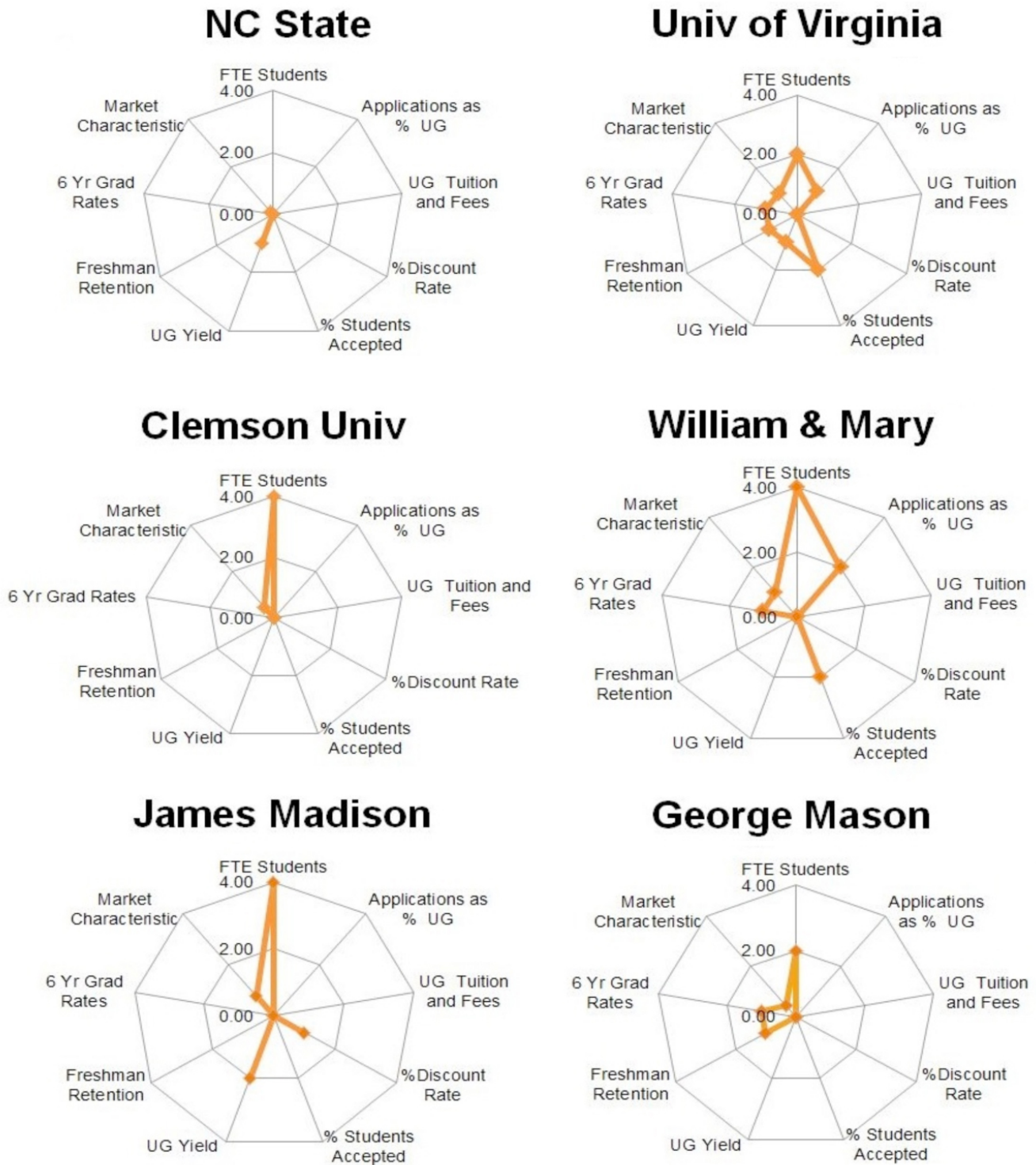


Figure 7 demonstrates how the results of nearest neighbor methods can be used to compare specific competitor universities to the focal institution. North Carolina State University, also a land-grant university, is most similar to the focal institution while the College of William & Mary, a liberal arts university, is least similar—although the two institutions are fairly similar in terms of six-year graduation rates and market characteristics.

Figure 7 Competitors to a Southeastern Land-Grant Research University



Figures 5–7 are only samples of the visualizations and charts that can describe the proximity of the focal institution to other institutions. As noted, the charts show the focal institution as the center of the comparisons. Plotting the distribution of institutional scores on key metrics relative to the focal institution is helpful to strategic planners who may need to visualize the relative performance of their institution compared to other similar institutions.



## SUMMARY AND LESSONS LEARNED

The preceding discussion describes a nearest neighbor methodology for forming reference groups to support institutional benchmarking. It demonstrates this methodology using publicly-available IPEDS data for a major southeastern land-grant research university.

There are important points that follow from experience with the reference group formation process. First, in today's higher education environment, institutions are not faced with the choice of having reference groups but rather with how to choose those reference groups. The reference group formation process is made more or less complex by the general public's fascination with the institutional rankings that are provided through numerous mechanisms. For example, vanity ratings using various criteria to group, compare, and rank institutions are published by the popular press and read by large numbers of citizens. The U.S. government also serves these citizens by grouping institutions through its College Navigator website, as do organizations such as the Institute for College Access and Success and NACUBO. All are providing mechanisms to facilitate comparisons between a focal institution and other "similar" institutions. Therefore, the question is not whether "your institution" will be compared, but to whom and how it will be compared.

Second, reference groups are now considered to be necessary for strategic planning and institutional assessment. Using nearest neighbor methods has utility in that their use supports the quantitative objectivity of analytics, is compatible with use of available national databases, and utilizes the judgmental expertise of key stakeholders.

Third, the role of stakeholders in setting parameters for reference group formation is important. Bringing stakeholders into discussions about the selection of reference institutions sends a message that their participation is critical to success. Reference group formation is not simply an analytic process but one that requires the judgment and input of knowledgeable stakeholders. Analytical methodologies can support judgment by creating visualizations that inform important stakeholders and support their understanding of why certain institutions may be appropriate for inclusion in the reference group.

*Reference group formation is not simply an analytic process but one that requires the judgment and input of knowledgeable stakeholders.*

Fourth, the availability of sufficient data is critical to success in forming reference groups. In this study, IPEDS data were used and shown to be appropriate for use in reference group formation. However, information made available on the IPEDS website reveals that data elements and data definitions sometimes change from year to year. Once downloaded, the data and all aspects of the process must be reviewed by knowledgeable analysts to determine that they are appropriate for use in the process of reference group formation.

Fifth, spreadsheets are used for the output of the nearest neighbor analytic process. Because the spreadsheets are large and complex, reviewing the outcomes with a stakeholder is much easier if the individual has a working knowledge of the software program (in this case, Excel). When this is not the case, a knowledgeable person who can effectively communicate the results to the stakeholder needs to be present. The nature of the conversation with the stakeholder is critical to successfully moving the reference group formation process forward.

Finally, the flexibility of nearest neighbor methods is best demonstrated by guiding relevant stakeholders through the analytical process. This is a critical step in their understanding of how the model works and how weights for each measure can be used to customize the model for a particular purpose. Going through this process also gives the stakeholder a greater appreciation for how the model can be used and more confidence in the appropriateness of the resulting comparison group. Giving the stakeholder the capacity to test different scenarios and to work with other campus leaders enhances the model's usefulness to help manage change.

## REFERENCES

- Bender, B. E., and J. H. Schuh, eds. 2002. *Using Benchmarking to Inform Practice in Higher Education*. New Directions for Higher Education, no. 118. San Francisco: Jossey-Bass.
- Blankmeyer, E., J. P. LeSage, J. R. Stutzman, K. J. Knox, and R. K. Pace. 2010. Peer-Group Dependence in Salary Benchmarking: A Statistical Model. *Managerial and Decision Economics* 32 (2): 91–104.
- Borden, V. M. H. 2005. Identifying and Analyzing Group Differences. In *Application of Intermediate/Advanced Statistics in Institutional Research*, ed. M. A. Coughlin, 132–68. Tallahassee, FL: Association for Institutional Research.
- Brinkman, P. T. 1987. Effective Institutional Comparisons. In *Conducting Institutional Comparisons*, ed. P. T. Brinkman, 103–108. New Directions for Institutional Research, no. 53. San Francisco: Jossey-Bass.
- Brinkman, P. T., and D. J. Teeter. 1987. Methods for Selecting Comparison Groups. In *Conducting Institutional Comparisons*, ed. P. T. Brinkman, 5–23. New Directions for Institutional Research, no. 53. San Francisco: Jossey-Bass.
- Gaither, G., B. P. Nedwek, and J. E. Neal. 1994. *Measuring Up: The Promises and Pitfalls of Performance Indicators in Higher Education*. ASHE-ERIC Higher Education Report, no. 5. San Francisco: Jossey-Bass.
- Health Information and Quality Authority. 2011. *International Review of Data Quality*. Dublin: Health Information and Quality Authority.
- Howard, R. D., G. W. McLaughlin, and J. S. McLaughlin. 1989. Bridging the Gap between the Data Base and User in a Distributed Environment. *CAUSE/EFFECT* 12 (2): 19–25.
- Kerschbaum, F. 2008. Building a Privacy-Preserving Benchmarking Enterprise System. *Enterprise Information Systems*, February: 1–15.
- Korb, R. 1982. *Clusters of Colleges and Universities: An Empirically Determined System*. Washington, DC: National Center for Education Statistics.
- Lang, D. W. 2000. Similarities and Differences: Measuring Diversity and Selecting Peers in Higher Education. *Higher Education* 39 (1): 93–129.
- McLaughlin, G. W., and J. S. McLaughlin. 2007. *The Information Mosaic: Strategic Decision Making for Universities and Colleges*. Washington, DC: Association of Governing Boards of Universities and Colleges.
- McLaughlin, J., D. Pavelka, and G. McLaughlin. 2005. Assessing the Integrity of Web Sites Providing Data and Information on Corporate Behavior. *Journal of Education for Business* 80 (6): 333–37.
- Nisbet, R., J. Elder, and G. Miner. 2009. *Handbook of Statistical Analysis and Data Mining Applications*. Burlington, MA: Elsevier.
- Pike, G. R., and G. D. Kuh. 2005. A Typology of Student Engagement for American Colleges and Universities. *Research in Higher Education* 46 (2): 185–209.

- Qayoumi, M. H. 2012. *Benchmarking and Organizational Change*. 2nd ed. Alexandria, VA: APPA.
- Reiss, E., S. Archer, R. Armacost, Y. Sun, and Y. (H.) Fu. 2010. Using SAS® PROC CLUSTER to Determine University Benchmarking Peers. Presentation given at the SESUG 2010 conference, Savannah, GA, September 27. Retrieved January 31, 2013, from the World Wide Web: [http://uaps.ucf.edu/doc/SESUG\\_Benchmarking\\_2010.pdf](http://uaps.ucf.edu/doc/SESUG_Benchmarking_2010.pdf).
- Secor, R. 2002. Penn State Joins the Big Ten and Learns to Benchmark. In *Using Benchmarking to Inform Practice in Higher Education*, ed. B. E. Bender and J. H. Schuh, 65–77. New Directions for Higher Education, no. 118. San Francisco: Jossey-Bass.
- StatSoft.com. n.d. How to Group Objects into Similar Categories, Cluster Analysis. Retrieved January 31, 2013, from the World Wide Web: [www.statsoft.com/textbook/cluster-analysis/?button=1](http://www.statsoft.com/textbook/cluster-analysis/?button=1).
- Teeter, D. J., and P. T. Brinkman. 1987. Peer Institutional Studies/Institutional Comparisons. In *A Primer on Institutional Research*, ed. J. Muffo and G. McLaughlin, 89–100. Tallahassee, FL: Association for Institutional Research.
- . 1992. Peer Institutions. In *Primer for Institutional Research*, ed. M. A. Whiteley, J. D. Porter, and R. H. Fenske, 63–72. Tallahassee, FL: Association for Institutional Research.
- Teeter, D. J., and M. E. Christal. 1987. Establishing Peer Groups: A Comparison of Methodologies. *Planning for Higher Education* 15 (2): 8–17.
- Terenzini, P. T., L. Hartmark, W. G. Lorang, Jr., and R. C. Shirley. 1980. A Conceptual and Methodological Approach to the Identification of Peer Institutions. *Research in Higher Education* 12 (4): 347–64.
- Townsley, M. T. 2002. *The Small College Guide to Financial Health: Beating the Odds*. Washington, DC: National Association of College and University Business Officers.
- Trainer, J. F. 2008. The Role of Institutional Research in Conducting Comparative Analysis of Peers. In *Institutional Research: More Than Just Data*, ed. D. G. Terkla, 21–30. New Directions for Higher Education, no. 141. San Francisco: Jossey-Bass.
- Wang, R. Y., and D. M. Strong. 1996. Beyond Accuracy: What Data Quality Means to Data Consumers. *Journal of Management Information Systems* 12 (4): 5–34.
- Weeks, S. F., D. Puckett, and R. Daron. 2000. Developing Peer Groups for the Oregon University System: From Politics to Analysis (and Back). *Research in Higher Education* 41 (1): 1–20.
- Xu, J. 2008. Using the IPEDS Peer Analysis System in Peer Group Selection. *Association For Institutional Research Professional File*, no. 110. Retrieved January 31, 2013, from the World Wide Web: <http://airweb3.org/airpubs/110.pdf>.

---

## DATA SOURCES FOR REFERENCE GROUP FORMATION

- ACT: [www.act.org](http://www.act.org)
- American Association of University Professors (AAUP): [www.aaup.org/AAUP/comm/rep/Z/](http://www.aaup.org/AAUP/comm/rep/Z/)
- Carnegie Foundation for the Advancement of Teaching: <http://classifications.carnegiefoundation.org/>
- College Results Online: [www.collegeresults.org/](http://www.collegeresults.org/)
- Consortium for Student Retention Data Exchange (CSRDE): <http://csrde.ou.edu/web/index.html>
- College and University Professional Association for Human Resources (CUPR-HR): [www.cupahr.org/](http://www.cupahr.org/)
- College Navigator, National Center for Education Statistics: <http://nces.ed.gov/collegenavigator/>
- Delaware Study of Instructional Costs and Productivity: [www.udel.edu/IR/cost/CAR/index.html](http://www.udel.edu/IR/cost/CAR/index.html)
- GuideStar: [www.guidestar.org/](http://www.guidestar.org/)
- Institute for College Access and Success: <http://ticas.org/>
- Integrated Postsecondary Education Data System, National Center for Education Statistics: <http://nces.ed.gov/ipeds/>
- National Association of College and University Business Officers (NACUBO): [www.nacubo.org/Research/NACUBO\\_Benchmarking\\_Tool.html](http://www.nacubo.org/Research/NACUBO_Benchmarking_Tool.html)
- National Student Clearinghouse: [www.studentclearinghouse.org](http://www.studentclearinghouse.org)
- National Survey of Student Engagement (NSSE): <http://nsse.iub.edu/>
- *U.S. News & World Report*: [www.usnews.com](http://www.usnews.com)

## AUTHOR BIOGRAPHIES

**Gerald W. McLaughlin** is an associate vice president for enrollment management and marketing at DePaul University and was formerly director of institutional research and planning analysis at Virginia Tech. He has been active in SCUP, EDUCAUSE, AIR, SAIR, and other national and international professional associations and was the former editor of *The AIR Professional File* and *IR Applications*. His interests include data management, methodology, benchmarking, planning, and strategic management.

**Josetta S. McLaughlin** is an associate professor of management at Roosevelt University where she teaches courses in strategic management, sustainability, and corporate social responsibility. She has been active in AOM, IABS, AIR, SAIR, CSRDE, SCUP, and other professional associations. Her research interests include strategic planning in higher education, social reporting, sustainability, and corporate corruption.

**Richard D. Howard** is retired from the University of Minnesota where he served as director of institutional research and professor of higher education. Over the past 35 years he has been active in AIR, SAIR, CAUSE, SCUP, and CSRDE. His professional interests include the theory and practice of institutional research, analytic methodologies, institutional benchmarking, and data management.