WHITE PAPER SERIES

Selecting Peer Institutions Using Cluster Analysis - Summer, 2021

Decision Support and Institutional Research

Austin Peay State University

About the Author

Dr. Andrew L. Luna, is Executive Director of Decision Support and Institutional Research. He has served over 36 years in higher education, with 25 of those years in institutional research. He has published research studies on many topics including salary studies, assessment, market research, and quality improvement. Dr. Luna received his Ph.D. and M.A. degrees in higher education administration and his M.A. and B.A. degrees in journalism, all from the University of Alabama.

Table of Contents

EXECUTIVE SUMMARY

In 2018 the Office of the Provost and Vice President of Academic Affairs charged Decision Support and Institutional Research (DSIR) with the task of creating a more scientific and reliable method for selecting APSU's peers. The method used is referred to as cluster analysis, which is defined as an exploratory data analysis technique for classifying and organizing data into meaningful clusters, groups, or taxonomies by maximizing the similarity between observations within each cluster.

Since the publication of that report, APSU began to offer its first doctoral program and has seen a change in the upper level administration. To support these changes, another peer analysis report was created in order to align APSU with updated peer institutions and replace the 2018 report.

The study established a base using all master-level (large programs) and doctoral/professional institutions with enrollments between 6,000 and 16,000 that were not Historically Black Colleges and Universities, land grant, or tribal institutions (n=113). Based upon the analysis, DSIR recommends including institutions from two closely-related clusters and that are also accredited by the Southern Association of Colleges and Schools Commission on Colleges (SACSCOC):

1. College of Charleston (South Carolina)
2. Columbus State University (Georgia)
3. Georgia College and State University
4. Jacksonville State University (Alabama)
5. McNeese University (Louisiana)
6. Morehead State University (Kentucky)
7. Northwestern State University of Louisiana
8. The University of Texas at Tyler
9. University of Houston - Clear Lake (Texas)
10. University of Louisiana at Monroe
11. University of North Alabama
12. Valdosta State University (Georgia)

This peer comparator study was reviewed and approved by the President's Senior Leadership Team on September 9, 2021.

INTRODUCTION

Within the current state of higher education, colleges and universities must strive to be competitive in both the quality of education they offer as well as the cost of attendance. At the same time, higher education is being held more accountable by federal and state governments, as well as by the communities they serve. This accountability varies broadly by legislative bodies, governors' offices, faculty committees, federal mandates, students and other constituencies. Therefore, the use of comparator institutions as a reference point within higher education has become common practice.

The use of peer comparator institutions allows administrators to compare both the quality and quantity of academic programs and delivery methods, as well as institutional expenditures and revenues. Comparisons like these allow for more focused strategic and long-range planning strategies in order to meet goals and objectives.

When identifying peers, it is important to understand the focus for the comparison group, as more than one set of peer groups may be utilized by an institution. There are various kinds of peers, such as:

- **Comparable:** Similar institutional level (two-year vs. four-year), control (e.g. private not-for-profit vs. public) and enrollment profile characteristics.
- **Aspirational:** Institutions with similar institutional characteristics yet are significantly different in several key performance indicators, such as significantly higher graduation rates or endowments.
- **Competitors:** Based on cross applications, institutions may have different institutional characteristics, yet a significant percentage of the institution's applicants choose to attend another institution.
- **Consortium:** Institutions belonging to a consortium for a common purpose and/or to share data

The purpose of this study is to identify those institutions that are comparable to APSU. These peer institutions share the same basic Carnegie Classification (e.g. Master's Institution [Large Program] and Doctoral/Professional), in addition to similar graduation rates and enrollment mix.

"The purpose of this study is to identify those institutions that are comparable to APSU. These peer institutions share the same basic Carnegie Classification (e.g. Master's Institution [Large Program] and Doctoral/Professional), in addition to similar graduation rates and enrollment mix."

Since 2013, APSU created its list of peer institutions for use as comparators in the Integrated Post Secondary Educational Data System (IPEDS) Institutional Feedback report. This peer list (below) was created from institutions located only in Tennessee:

1. East Tennessee State University (Johnson City, TN)
2. Middle Tennessee State University (Murfreesboro, TN)
3. Tennessee State University (Nashville, TN)
4. Tennessee Technological University (Cookeville, TN)
5. The University of Tennessee-Chattanooga (Chattanooga, TN)
6. The University of Tennessee-Martin (Martin, TN)
7. University of Memphis (Memphis, TN)

APSU also worked with the National Survey of Student Engagement (NSSE) to develop several comparison groups for data analysis based on type, size, location, and Carnegie Classification. Selections for the NSSE comparison groups are limited by the options provided by NSSE, as well as the Tennessee Higher Education Commission's requirements for Quality Assurance Funding reporting.

Additionally, APSU partnered with the College and University Professional Association (CUPA) to conduct various salary studies. As part of this process, CUPA developed a set of peer institutions as a base for salary comparisons. This list has not changed since 2018. While many institutions within the CUPA list are comparable to APSU, the list of 68 institutions (**Appendix III**) is also broad in that it includes research institutions, institutions with over 15,000 enrollment, as well as Historically Black Colleges and Universities (HBCU).

Sensing a need to create a peer list that had closer similarities to APSU, the Provost and Vice President for Academic Affairs' Office charged DSIR in 2018 with the task of creating a more scientific and reliable method for selecting APSU's peers.

With significant changes occurring within APSU such as the offering of a doctoral program as well as changes in the upper-level administration, there existed a need to update and replace the 2018 analysis. However, the currently methodology of cluster analysis is similar to what was used in 2018.

The process of utilizing statistical methodologies in the identification of peer institutions began more than 20 years ago (Teren-

"The process of utilizing statistical methodologies in the identification of peer institutions began more than 20 years ago."

zini, et al., 1980; Teeter & Brinkman, 1987; and McLaughlin & McLaughlin, 2007). The overall goal during this time has been to identify appropriate methods for comparing the performance of a reference institution relative to a group of similar institutions, and to make goal and outcome decisions concerning the reference institution based on the performance of the comparator institutions.

While the use of statistical methodologies supports scientific objectivity, their complexity often makes them difficult to understand by the end user. Other studies have also indicated that these types of methodologies inherently contain statistical error due to the additive and multiplicative attributes of the procedures used (McLaughlin & McLaughlin, 2007). It is, therefore, recommended that the institution not rely solely on the outcome of a statistical peer analysis. Rather, the data from the analysis should be used in conjunction with other knowledge gained.

"...cluster analysis, [is] defined as an exploratory data analysis technique for classifying and organizing data into meaningful cluster, groups, or taxonomies..."

This study used cluster analysis, which is defined as an exploratory data analysis technique for classifying and organizing data into meaningful clusters, groups, or taxonomies by maximizing the similarity between observations within each cluster. The purpose of cluster analysis is to discover a system of organizing observations into groups where members of the groups share properties in common. The goal of this analysis, therefore, is to sort variables into groups or clusters so that the degree of association or relationship is strong between members of the same cluster and weaker between members of different clusters.

The appropriate cluster algorithm and parameter settings depend on the individual data set and intended use of the results. Furthermore, cluster analysis is an iterative process of knowledge discovery and optimization to modify data processing and model parameters until the result achieves both the preferred as well as appropriate properties.

The choice of methods used for cluster analysis depends on the size of the data set as well as the types of variables used. In this study, hierarchical clustering is more appropriate because the data set is small. The steps in obtaining and preparing the data for cluster analysis are as follows:

- Screen institutions to determine what type and size of institution will be used in the analysis

- Choose variables to download from IPEDS that will be used in the analysis
- Standardize all quantifiable variables that will be used in the analysis
- Run the cluster analysis procedure
- Determine the fit and reliability of the model
- Identify those institutions that are within the same cluster as APSU

IPEDS INITIAL INSTITUTIONAL SCREENING

To start the process of determining institutional peers, an initial reference group was established. Larger research institutions, two-year colleges, and specialty institutions with a significantly different role, scope, and mission than APSU were screened out. Additionally, the updated study includes professional doctoral institutions since APSU's doctoral program was first offered after the publication of the 2018 study. A list of institutions was generated through the IPEDS system by choosing only public 4-year institutions with a Carnegie Classification of Masters - Larger programs, doctoral/professional programs, and total enrollment between 6,000 and 16,000.

"Larger research institutions, two-year colleges, and specialty institutions with a significantly different role, scope, and mission were screened out."

From these criteria, a total of 113 institutions were included for the cluster analysis. This result was significantly higher than the 73 institutions generated during the 2018 study. A listing of all institutions used in this study can be found in **Appendix IV**. From these institutions, specific variables were chosen to be used in the cluster analysis procedure.

Choosing Variables to Use in the Analysis

Once the initial 113 institutions were selected, a total of 12 selected variables were downloaded from the IPEDS Data Center for each institution. These variables were selected by DSIR following an extensive literature review process on what key variables are factors in determining institutional role, scope, and mission. The variables selected are listed below:

1. Undergraduate enrollment for latest fall semester
2. Graduate enrollment for latest fall semester
3. FTE for latest academic year
4. Six-year graduation rate based on the IPEDS defined freshman cohort
5. Total operational revenues
6. Tuition and fees as a percent of operational revenues
7. State appropriations as a percent of operational revenues
8. Total expenditures
9. Instructional costs as a percent of expenditures
10. Endowment assets per FTE
11. In-state tuition and fees on-campus
12. Out-of-state tuition and fees on-campus

Once the variables for the study were pulled, it was determined that data values of the 12 variables were present in each of the 113 institutions. Therefore, none of the original 113 institutions were removed based upon insufficient data.

Given that the raw data pulled from IPEDS for this analysis significantly vary, all of the variables were then standardized for use in the analysis. It should be noted that using variables without standardizing them can give those variables with larger values and ranges greater importance in the analysis. Standardizing the variables remedies this issue. The standardization used in this study is reviewed in **Appendix I** of this report.

"Once the variables for the study were pulled, it was determined thatdata values of the 12 variables were preswent in each of the 113 institutions"

RUNNING THE CLUSTER ANALYSIS PROCEDURE

The objective of cluster analysis is to group observations of interest into clusters so that those observations within each group are similar inside the group while each group stands apart from each other.

Take, for example, a group of people who are inside a stadium. As one large group, there exists a lot of variability and difference. If the larger group was parsed out by certain key variables or attributes, those people who were married, earned high income, and had multiple children would fall into one cluster or group while individuals who had lower incomes, were single, and had no children would be placed into another group. In this example, individuals having similar attributes would be in the same cluster while those who were different would fall into another cluster.

While there are many ways to run a cluster analysis, there are two basic fundamental methods of hierarchical (systematic) analysis. The first method involves forming as many groups as there are observations and the systematically merging observations in order to reduce the number of groups. This method is call agglomerative. The second basic method is called divisive in that it groups all observations into one cluster and then separates the observations into like groups. This study uses the former, agglomerative, method.

It is important, however, to note that whatever method is used, the risk of under or over specifying the model may occur. For example, if there are 80 observations, there is clearly little benefit in grouping all 80 into one group or, likewise, to place each observation into one of 80 clusters.

The mechanics of running a cluster analysis involves determining distance of each variable or attribute within an observation and grouping similar distances together. For a more detailed explanation of cluster analysis, please refer to **Appendix II** of this report.

"While there are numerous ways in which clusters may be formed, hierarchical clustering is one of the most straightforward methods."

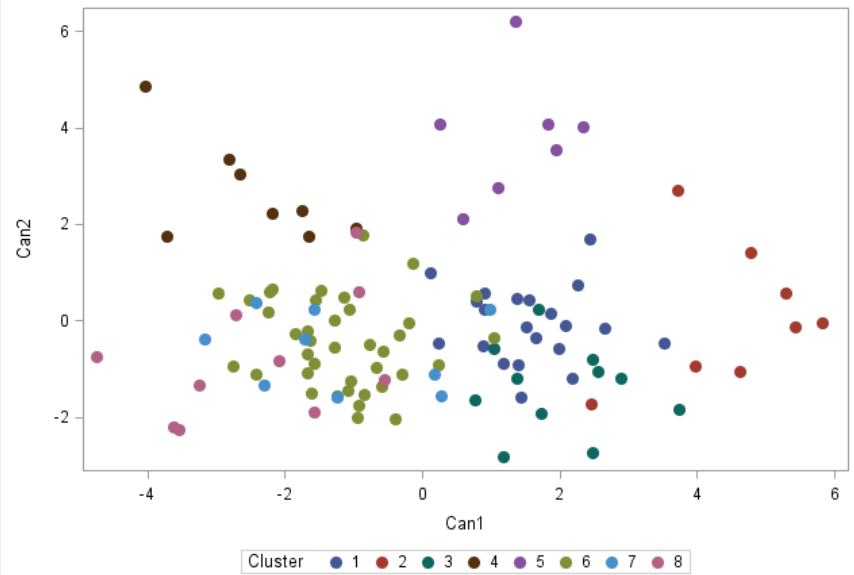
DETERMINING FIT AND RELIABILITY OF MODEL

After the cluster analysis procedure determined that the 113 institutions could be reduced into eight unique clusters, a canonical discriminant analysis was run to create grouped variables for use in a scatter plot in order to determine where each of the clusters fall. Canonical discriminant analysis is used to find a linear combination of features which characterizes or separates two or more classes of objects or events. The resulting combination may be used as a linear classifier or, more commonly, for dimensionality reduction before later classification.

In essence, the canonical discriminant analysis determines distances between one or more quantitative variables and then determines the relationship between the quantitative variables and a set of classification variables to place observations into clusters so that every observation belongs to one and only one cluster.

The first canonical correlation is the maximum correlation that can be obtained between a linear combination of one set of variables and a linear combination of another set of variables. The second canonical correlation is the maximum correlation that can be obtained between linear combinations of the two sets of variables subject to the constraint that these second linear combinations are orthogonal (independent/uncorrelated) to the first linear combinations. The second canonical variable provides the greatest difference between group means while being uncorrelated with the first canonical variable.

Table 1: Scatter Plot of Clusters
Standardized Variables



Within this study, the R^2 value at .47 was significant and the first canonical correlation indicated a .90 which was a higher value than the second cononical correlation, so plotting the first canonical correlation should give a good

indication where the clusters fall and how closely related they are to each other. Following the FASTCLUS procedure in SAS, the first canonical variable was plotted against the second canonical variable. Together, these variables indicate where the various clusters reside, how widely distributed they are, and how close they are to each other. As can be seen in **Table 1**, all of the clusters are distinct, albeit some are close together with clusters 2 (mauve), 6 (green), and 7 (lt. blue) overlapping in some places. APSU is within cluster 6.

In determining which cluster the home institution resides, it is also important to note a couple of other pieces of information from the scatter plot. First, some clusters are very close to the home institution's cluster and, in some cases, may actually be intertwined within the home cluster. If there was a desire to remove institutions from the home cluster due to geographic reasons, other institutions could be used from clusters that are close to the home cluster.

It should be noted that in determining the final peer group, some of the institutions within the home cluster that reside outside of the southeast may be replaced with southeastern institutions residing in clusters close to or intertwined within the home cluster.

Another piece of information to consider while observing the scatter plot is to see the relationships of the other clusters to the home cluster. For example, in observing clusters ranked above the home cluster, the institution may want to look at institutions within these clusters as possible aspirational peers.

"It should be noted that in determining the final peer group, some of the institutions within the home cluster that reside outside of the southeast may be replaced with southeastern institutions residing in clusters close to or intertwined within the home cluster."

RESULTS

The FASTCLUS procedure within SAS indicated that the 113 institutions would best be divided into eight clusters with APSU residing in cluster 6. The results also indicate that cluster 7 was close to the APSU cluster. Those institutions within cluster 6 were:

1. Arkansas Tech University
2. Bloomsburg University of Pennsylvania
3. Bridgewater State University (Massachusetts)
4. California University of Pennsylvania
5. Columbus State University (Georgia)
6. East Stroudsburg University of Pennsylvania
7. Jacksonville State University (Alabama)
8. Kutztown University of Pennsylvania
9. McNeese State University (Louisiana)
10. Morehead State University (Kentucky)
11. New Jersey City University
12. Northeastern State University (Oklahoma)
13. Northern Michigan University
14. Northwest Missouri State University
15. Northwestern State University of Louisiana
16. Purdue University Fort Wayne (Indiana)
17. Purdue University Northwest (Indiana)
18. Rhode Island College
19. SUNY College at Brockport
20. SUNY College at Oswego
21. Saginaw Valley State University (Michigan)
22. Saint Cloud State University
23. Salem State University (Massachusetts)
24. Southeast Missouri State University
25. Southern Utah University
26. State University of New York at New Paltz
27. The University of Texas at Tyler
28. University of Central Arkansas
29. University of Houston-Clear Lake
30. University of Louisiana at Monroe
31. University of North Alabama
32. University of Southern Maine
33. University of Wisconsin-Green Bay
34. University of Wisconsin-Platteville
35. University of Wisconsin-Stout
36. Valdosta State University (Georgia)
37. Western Illinois University

"With the cluster analysis indicating that cluster 7 was close to cluster 6, further investigations as to which institutions reside within this cluster should be done."

Within this cluster, 10 institutions reside in the same geographic/accreditation region as APSU with Columbus State University, Jacksonville State University, McNeese State University, Morehead State University, Northwestern State University of Louisiana, The University of Texas at Tyler, University of Houston-Clear Lake, University of Louisiana at Monroe, and the University of North Alabama, and Valdosta State University sharing the same regional accreditation as APSU.

The cluster analysis indicated that cluster 7 was also close to APSU's home cluster, suggesting that further investigation should be conducted to determine if any of these institutions should be added to APSU's peer list or used to replace cluster 6 institutions outside the accreditation region. According to the analysis, those institutions within cluster 7 were:

1. College of Charleston (South Carolina)
2. Georgia College & State University
3. Millersville University of Pennsylvania
4. SUNY Cortland
5. SUNY Oneota
6. Salsbury University
7. Sonoma State University
8. Stockton University
9. The College of New Jesery

Within this cluster, two of the institutions share the same geographic/accreditation region (i.e. College of Charleston and Georgia College & State University). Therefore, some of these institutions could possibly be used to replace other institutions within the home cluster (cluster 6) that are not in the same accreditation region as APSU.

Table 2: APSU Values Compared to Primary and Secondary Clusters

Variables Used in Study	APSU Value	Primary Cluster Mean	Secondary Cluster Mean
Undergraduate enrollment for latest fall semester	9,971	7,689.28	7,386.00
Graduate enrollment for latest fall semester	1,077	1,297.10	834.00
FTE for latest academic year	9,115	7,639.28	7,988.33
Six-year graduation rate based on the IPEDS defined freshman cohort	41	49.38	69.89
Total operational revenues	77,864,054	89,657,116.35	126,375,952
Tuition and fee as percent of operational revenues	69.17	60.94	56.37
State appropriations as a percent of operational expenditures	63.27	51.54	53.52
Total expenditures	171,276,244	359,465,996	212,156,505
Instructional costs as a percent of expenditures	43.39	38.77	34.06
Endowment Assets per FTE	2,467.49	4,906.90	6,395.22
In-state tuition and fees on-campus	27,712.00	23,724.78	29,181.44
Out-of-state tuition and fees on-campus	33,256.00	32,499.13	41,461.004

The data listed in **Table 2** includes all of the twelve variables used in the study, APSU's value for each of these variables, and the mean values of each variable for cluster 6 (Primary), and cluster 7 (Secondary). From these data, it is clear that many of APSU's values more closely align with the means

of cluster 6 than they do with cluster 7. The exceptions are the undergraduate enrollment where APSU is higher than both clusters; graduate enrollment where APSU is closer to cluster 6, and six-year graduation rate where APSU is lower than both clusters, but closer to cluster 6. According to this study, none of the other Tennessee 4-year public institutions were present in either cluster 6 or 7.

The purpose of this study is to separate select institutions into similar clusters for use in determining comparator peers for APSU. The proximity of the cluster that the institutions are being compared should be considered as well as such factors as cost-of-living, non-traditional and international enrollment, location of major metropolitan areas close to the institution, and regional accrediting associations. These factors could significantly affect comparability within any model. Therefore, choosing institutions sharing the same regional accreditation as APSU was a major factor in DSIR's recommendation.

Recommendations

Based on the cluster analysis outcomes from this study, along with external factors such as cost-of-living and accreditation considerations, Decision Support and Institutional Research recommends 12 institutions that were within both clusters 6 and 7. The information within **Table 3** indicates the institution chosen, which cluster the institution was grouped, if the institution is in APSU's current list of peers, and if the institution was included within CUPA's list of peers for APSU.

Table 3: Recommended Comparator Peer Institutions

Institution	Cluster No.	2018 Peer	CUPA Peer
College of Charleston (South Carolina)	7	No	Yes
Columbus State University (Georgia)	6	Yes	Yes
Georgia College & Sate University	7	No	Yes
Jacksonville State University (Alabama)	6	Yes	Yes
McNeese State University (Louisiana)	6	No	Yes
Morehead State University (Kentucky)	6	Yes	Yes
Northwestern State University of Louisiana	6	No	Yes
The University of Texas at Tyler	6	Yes	Yes
University of Houston-Clear Lake (Texas)	6	Yes	Yes
University of Louisiana at Monroe	6	No	No
University of North Alabama	6	Yes	Yes
Valdosta State University (Georgia)	6	No	Yes

All of these institutions are within APSU's regional accreditation area. Furthermore, all but one of the selected institutions were also included in the CUPA peer list. Therefore, this new list includes both institutions from the original 2018 peer list and the CUPA peer list while only including public 4-year Masters (Larger Programs) and professional doctoral programs with enrollments between 6,000 and 16,000 students.

Aspirational Peers

In addition to determining which instituions can be comparable to APSU, the study also identified those clusters that are physically more advanced on the scatterplot. These institutions become possible aspirational comparators. These aspirational peers would be classified as institutions that have characteristics APSU would like to emulate in the future. By looking at the scatterplot in **Table 1**, it can be determined that clusters 4 and 8 could be potential aspirational peers. Therefore, **Table 4** provides a listing of potential aspirational peers, from which cluster they reside, and whether these institutions were in the previous 2018 study and or from the CUPA Peer list. From these, it is recommeded that the following institutions be considered aspirational peers because they share the same regional accreditation as APSU:

Table 4: Possible Aspirational Peer Institutions

Institution	Cluster No.	2018 Peer	CUPA Peer
Angelo State University (Texas)	4	No	No
Murray State University (Kentucky)	4	Yes	Yes
Pittsburg State University (Kansas)	4	No	No
The University of Tennessee-Chattanooga	4	Yes	Yes
University of Minnesota-Duluth	4	No	No
University of Southern Indiana	4	No	No
West Texas A & M University	4	No	Yes
Youngstown State University (Ohio)	4	No	No
California State University-Bakersfield	8	No	No
California State University-Channel Islands	8	No	No
California State University-Monterey Bay	8	No	No
California State University-Stanislaus	8	No	No
Humboldt State University (California)	8	No	No
SUNY Buffalo State (New York)	8	No	No
Texas A & M International University	8	Yes	Yes
The University of West Florida	8	No	Yes
University of North Carolina at Pembroke	8	Yes	Yes
Western Carolina University (North Carolina)	8	No	Yes

1. Murray State University
2. The University of Tennessee-Chattanooga
3. West Texas A & M University
4. Texas A & M International University
5. The University of West Florida
6. University of North Carolina at Pembroke
7. Western Carolina University

In addition to identifying aspirational peers, further analysis can be performed by using different variables from IPEDS. This allows the institution to more closely align itself with institutional peers based on specific variables. Furthermore, the institution could use the IPEDS initial screening to include a peer group that is confined to only one region of the US.

It is important to note that subsequent iterations of the cluster analysis can include variables that were not pulled from the IPEDS database. Clearly, the number of total variables can impact the reliability of the model given the relatively small number of institutions, and parsimony is preferred. However, the model does allow administrators to choose variables that more closely align with the institution's role, scope, and mission in order to create a more meaningful institutional peer group.

While cluster analysis is clearly an exploratory data analysis technique for classifying and organizing institutions into meaningful groups, the results of such analyses are not definitive and should be reviewed with other quantitative and qualitative criteria. These methods, however, can save time and resources as institutions seek to find peer institutions to match their benchmarking needs.

Standardizing all quantifiable variables used in the analysis

Many researchers have noted the importance of standardizing variables for multivariate analysis. Otherwise, variables measured at different scales may not contribute equally to the analysis. This practice holds true for cluster analysis. Because of the sensitivity of most cluster models, raw values used for the variables may significantly alter the outcomes.

For example, in selecting peer institutions, a variable that ranges between $5 million and $10 million will influence significantly and have more weight in the analysis than a variable that ranges between 20 and 50. Therefore, transforming the data to comparable scales can prevent this problem. Typical data standardization procedures equalize the range and/or data variability. In the case of this study, variable values were standardized using z-scores with a mean of zero and a standard deviation of 1.

The z-score is a very useful statistic because it allows researchers to calculate the probability of a score occurring within the normal distribution and it enables researchers to compare two scores from different normal distributions. The standard score converts scores in a normal distribution to z-scores using the following formula:

$$\frac{ii \quad -}{\quad\quad}$$

where represents an individual score or observation in a set of scores, represents the average of all individual scores or observations, and S represents the standard deviation of the scores or observations.

The z-score is synonymous to the standard deviation. A z-score of 2 is essentially 2 standard deviations above and below the mean. A z-score of 1.5 is 1.5 standard deviations above and below the mean. A z-score of 0 is equal to the mean of the distribution.

Z-scores exist on both sides of the mean. For example, 1 standard deviation below the mean is a z-score of -1 and a z-score of 2.2 can be 2.2 standard deviations above the mean. A z-score of -3 is 3 standard deviations below the mean. Put another way, the standard deviation and z-scores are just the average distance that individual values are from the mean.

Appendix I

"Clearly, the number of total variables used can impact the reliability of the model given the relatively small number of institution, and parsimony is preferred. However, the model does allow administrators to choose variables that more closely align with the institution's role, scope, and mission..."

Appendix II

Running the Cluster Analysis Procedure Using FASTCLUS within SAS

While there are numerous ways in which clusters may be formed, hierarchical clustering is one of the most straightforward methods. It can be either agglomerative or divisive. Agglomerative hierarchical clustering begins with each institution being a cluster unto itself. At successive steps, similar clusters are merged. The algorithm ends with all institutions in one, but useless, cluster. Divisive clustering starts with all institutions in one cluster and ends with each institution in its own cluster which, again, is not helpful. To find a good cluster solution, the researcher must look at the characteristics of the clusters at successive steps and decide when an interpretable solution is found that has a reasonable number of fairly homogeneous clusters.

This study used PROC FASTCLUS within SAS to determine the clusters. While the FASTCLUS procedure is intended for larger data sets, it can be used with smaller, although it can be sensitive to the order of the observations within the data set. This issue can be negated by standardizing the variables. PROC FASTCLUS also uses algorithms that place a large influence on variables with larger variance. Again, standardizing the variables before performing the analysis is highly recommended.

PROC FASTCLUS performs a disjoint cluster analysis on the basis of distances computed from one or more quantitative variables. The observations are divided into clusters so that every observation belongs to one cluster. By default, PROC FASTCLUS uses Euclidean distances, so the cluster centers are based on least squares estimation. The cluster centers are the means of the observations assigned to each cluster when the algorithm is run to complete convergence. PROC FASTCLUS is designed to find good clusters, not the best possible clusters, with only two or three iterations of the data set and changing the number of clusters requested. This procedure can be effective in detecting outliers which appear as clusters with only one institution.

To run the analysis a two-step process was used to determine the number of possible clusters. This process used the CLUSTER procedure within SAS in order to examine eigenvalues, differences, and proportions. According to **Table 5**, a large difference exists between the first (3.618) and second (1.931) eigenvalues, proportions go from .3015 to .1609, with the cumulative pro-

portion for the second eigenvalue equal to .4623. While this seems significant, a total of 113 institutions within only two clusters would be considerably under specified and the cumulative proportion indicates more clusters could be formed.

Upon further examination of the table, there exists a moderate change from eigenvalues eight (.5058) and nine (.2755), proportions go from .0422 to .0230, with the cumulative proportion for the ninth eigenvalue equal to .9738 which is not much different from the cumulative percentage of .9508 at eigenvalue eight. Further investigation revealed that clusters greater than eight would not contribute significantly to the model. Therefore, eight clusters were examined with results from PROC FASTCLUS.

Table 5: Eigenvalues of the Correlation Matrix

	Eigenvalue	Difference	Proportion	Cumulative
1	3.618	1.687	0.3015	0.3015
2	1.931	0.379	0.1609	0.4623
3	1.551	0.281	0.1293	0.5916
4	1.270	0.186	0.1058	0.6974
5	1.084	0.346	0.0903	0.7878
6	.7380	0.024	0.0615	0.8492
7	.7133	0.207	0.0594	0.9087
8	.5058	0.230	0.0422	0.9508
9	.2755	0.037	0.0230	0.9738
10	.2389	0.194	0.0199	0.9937
11	.0450	0.014	0.0037	0.9974
12	.3080		0.0026	1.0000

Running the FASTCLUS procedure on eight clusters generated a significant Pseudo F Statistic of 13.39 and an observed overall R-Squared value of .47. The multivariate statistics and F approximations were then computed to test the fit of the model and the Wilks' Lambda, Pillai's Trace, Hotelling-Lawley Trace, and Roy's Greatest Root all confirmed that the model was significant with eight clusters.

Appendix III

An Analysis of CUPA Peer Institutions

As part of APSU's process of analyzing on-campus salaries, the institution developed a comparator group of public institutions in 2002. In 2010, there was a revised compensation study and the peer group was slightly modified to its current 68 institutions. The comparator group has been used by CUPA for regularly-scheduled salary studies.

In examining the CUPA peer list for possible inclusion as APSU's office peer comparator list, it was noted that some of the institutions were not aligned with APSU's role, scope, and mission.

Specifically, out of the 68 institutions chosen by CUPA, 19 had the Carnegie Classification of Doctoral/Research, 16 institutions had enrollments over 16,000 students, seven were Historically Black Colleges and Universities (HBCU), and seven had enrollments under 6,000. Below are the institutions used within CUPA salary studies:

1. Alabama Agricultural and Mechanical University
2. Appalachian State University
3. Arkansas State University
4. Armstrong State University
5. Auburn University at Montgomery
6. Bowie State University
7. College of Charleston
8. Columbus State University
9. Delta State University
10. East Carolina University
11. Eastern Kentucky University
12. East Tennessee State University
13. Fayetteville State University
14. Frostburg State University
15. Georgia College & State University
16. Georgia Southern University
17. Grambling State University
18. Jackson State University
19. Jacksonville State University
20. James Madison University
21. Marshall University
22. McNeese State University
23. Middle Tennessee State University
24. Morehead State University
25. Morgan State University
26. Murray State University
27. North Carolina Agricultural and Technical State University
28. North Carolina Central University
29. Northeastern State University
30. Northern Kentucky University
31. Northwestern State University
32. Prairie View A & M University

33. Radford University
34. Salisbury University
35. Sam Houston State University
36. Southeastern Louisiana University
37. Stephen F. Austin State University
38. Tarleton State University
39. Tennessee State University
40. Tennessee Technological University
41. Texas A&M International University
42. Texas A&M University - Corpus Christi
43. Texas A&M University - Kingsville
44. Texas State University
45. The University of Memphis
46. The University of Texas At El Paso
47. Towson University
48. Troy University
49. University of Central Arkansas
50. University of Central Oklahoma
51. University of Houston - Clear Lake
52. University of Houston – Victoria
53. University of North Alabama
54. University of North Carolina at Charlotte
55. University of North Carolina at Pembroke
56. University of North Carolina - Wilmington
57. University of North Georgia
58. University of South Alabama
59. University of Tennessee at Chattanooga
60. University of Tennessee at Martin
61. University of Texas at Tyler
62. University of West Alabama
63. University of West Florida
64. Valdosta State University
65. Western Carolina University
66. Western Kentucky University
67. West Texas A & M University
68. Winthrop University

Appendix IV

Institutions Used in the Cluster Analysis Study

Institutions used in the study included all public 4-year institutions with a Carnegie Classification of Masters - Larger Programs, enrollment between 6,000 and 16,000, and were not classified as HBCU, Land Grant, or Tribal institutions. The following institutions were used in the study:

1. Angelo State University
2. Appalachian State University
3. Arkansas Tech University
4. Bloomsburg University of Pennsylvania
5. Bridgewater State University
6. California State University-Bakersfield
7. California State University-Channel Islands
8. California State University-Chico
9. California State University-Dominguez Hills
10. California State University-East Bay
11. California State University-Monterey Bay
12. California State University-San Marcos
13. California State University-Stanislaus
14. California University of Pennsylvania
15. Central Connecticut State University
16. Central Washington University
17. Coastal Carolina University
18. College of Charleston
19. College of Staten Island CUNY
20. Columbus State University
21. CUNY Bernard M Baruch College
22. CUNY John Jay College of Criminal Justice
23. CUNY Lehman College
24. CUNY Queens College
25. East Stroudsburg University of Pennsylvania
26. Eastern Kentucky University
27. Eastern Washington University
28. Ferris State University
29. Florida Gulf Coast University
30. Fort Hays State University
31. Georgia College & State University
32. Humboldt State University
33. Indiana State University
34. Indiana University of Pennsylvania-Main Campus
35. Jacksonville State University
36. Kean University
37. Kutztown University of Pennsylvania
38. Lamar University
39. McNeese State University
40. Millersville University of Pennsylvania
41. Minnesota State University-Mankato
42. Morehead State University
43. Murray State University
44. New Jersey City University
45. Northeastern State University

46. Northern Kentucky University
47. Northern Michigan University
48. Northwest Missouri State University
49. Northwestern State University of Louisiana
50. Pittsburg State University
51. Purdue University Fort Wayne
52. Purdue University Northwest
53. Radford University
54. Rhode Island College
55. Saginaw Valley State University
56. Saint Cloud State University
57. Salem State University
58. Salisbury University
59. Slippery Rock University of Pennsylvania
60. Sonoma State University
61. Southeast Missouri State University
62. Southeastern Louisiana University
63. Southern Connecticut State University
64. Southern Illinois University-Edwardsville
65. Southern Utah University
66. State University of New York at New Paltz
67. Stephen F Austin State University
68. Stockton University
69. SUNY Buffalo State
70. SUNY College at Brockport
71. SUNY College at Oswego
72. SUNY Cortland
73. SUNY Oneonta
74. Tarleton State University
75. Texas A & M International University
76. Texas A & M University-Commerce
77. Texas Woman's University
78. The College of New Jersey
79. The University of Tennessee-Chattanooga
80. The University of Texas at Tyler
81. The University of West Florida
82. Troy University
83. University of Alaska Anchorage
84. University of Central Arkansas
85. University of Central Missouri
86. University of Central Oklahoma
87. University of Houston-Clear Lake
88. University of Louisiana at Monroe
89. University of Minnesota-Duluth
90. University of North Alabama
91. University of North Carolina at Pembroke
92. University of North Florida
93. University of North Georgia
94. University of Northern Iowa
95. University of Southern Indiana
96. University of Southern Maine
97. University of West Georgia
98. University of Wisconsin-Eau Claire
99. University of Wisconsin-Green Bay
100. University of Wisconsin-La Crosse

101. University of Wisconsin-Oshkosh
102. University of Wisconsin-Platteville
103. University of Wisconsin-Stout
104. University of Wisconsin-Whitewater
105. Valdosta State University
106. West Chester University of Pennsylvania
107. West Texas A & M University
108. Western Carolina University
109. Western Illinois University
110. Western Kentucky University
111. Western Washington University
112. William Paterson University of New Jersey
113. Youngstown State University

References

McLaughlin, G.W. and McLaughlin, J.S. (2007). The information mosaic: strategic decision making for universities and colleges. AGB: Washington, DC.

Teeter, D. J and Brinkman, P.T. (1987). Peer institutional studies/institutional comparisons. Primer for Institutional Research, J. Muffo and G. McLaughlin (eds), Association for Institutional Research: Tallahassee.

Terenzini, P. T., Hartmark, L., Lorang, W. G., & Shirley, R. C. (1980). A conceptual and methodological approach to the identification of peer institutions. Research in Higher Education, 12, 347-364.