NATIONAL CENTER FOR EDUCATION STATISTICS

Methodology Report   **June 2001**

Postsecondary Education Descriptive Analysis Reports

A Classification System for 2-Year Postsecondary Institutions

Ronald A. Phipps
Jessica M. Shedd
Jamie P. Merisotis
The Institute for Higher Education Policy

C. Dennis Carroll
Project Officer
National Center for Education Statistics

U.S. Department of Education
Rod Paige
Secretary

National Center for Education Statistics
Gary W. Phillips
Acting Commissioner

The National Center for Education Statistics (NCES) is the primary federal entity for collecting, analyzing, and reporting data related to education in the United States and other nations. It fulfills a congressional mandate to collect, collate, analyze, and report full and complete statistics on the condition of education in the United States; conduct and publish reports and specialized analyses of the meaning and significance of such statistics; assist state and local education agencies in improving their statistical systems; and review and report on education activities in foreign countries.

NCES activities are designed to address high priority education data needs; provide consistent, reliable, complete, and accurate indicators of education status and trends; and report timely, useful, and high quality data to the U.S. Department of Education, the Congress, the states, other education policymakers, practitioners, data users, and the general public.

We strive to make our products available in a variety of formats and in language that is appropriate to a variety of audiences. You, as our customer, are the best judge of our success in communicating information effectively. If you have any comments or suggestions about this or any other NCES product or report, we would like to hear from you. Please direct your comments to:

> National Center for Education Statistics
> Office of Educational Research and Improvement
> U.S. Department of Education
> 1990 K Street NW
> Washington, DC 20006-5574

June 2001

The NCES World Wide Web Home Page is: ***http://nces.ed.gov***
The NCES World Wide Web Electronic Catalog is: ***http://nces.ed.gov/pubsearch/index.asp***

Suggested Citation
U.S. Department of Education. National Center for Education Statistics. *A Classification System for 2-Year Postsecondary Institutions,* NCES 2001–167, by Ronald A. Phipps, Jessica M. Shedd, and Jamie P. Merisotis. Project Officer: C. Dennis Carroll. Washington, DC: 2001.

For ordering information on this report, write:
U.S. Department of Education
ED Pubs
P.O. Box 1398
Jessup, MD 20794-1398

or call toll free 1-877-4ED-PUBS or go to the Internet: ***http://www.ed.gov/pubs/edpubs.html***

Contact:
Aurora D'Amico
(202) 502-7334

Executive Summary

Background

A well-known classification system for higher education in the United States was developed by the Carnegie Commission on Higher Education. Originally published in 1973—and refined and updated in 1976, 1987, 1994, and 2000—the Carnegie classification system divides colleges and universities into categories based on their degree-granting activity (doctoral, master's, baccalaureate, and associate's). In addition, the Carnegie classification system identifies myriad "specialized institutions," such as theological seminaries, schools of law, teachers colleges, and medical schools.

A limitation of the Carnegie classification system is how 2-year institutions are categorized. While there are six distinct subcategories for 4-year institutions, the 2000 Carnegie classification system places the 1,669 institutions that offer associate of arts degrees or certificates and—with few exceptions—offer no baccalaureate degrees into a single grouping, Associate's Colleges. This single grouping of institutions accounts for more institutions than Doctoral/Research Universities, Master's Colleges and Universities, and Baccalaureate Colleges combined (1,478).[1] Furthermore, the Carnegie classification includes only accredited and degree-granting institutions, excluding over 700 of the 2,427 2-year institutions with participation agreements for Title IV student aid

funds, as found in the 1997–98 Institutional Characteristics survey of the Integrated Postsecondary Education Data System (IPEDS) database.

About half of all students in postsecondary education are enrolled in 2-year institutions.[2] Placing all 2-year institutions into one category, thereby masking their differences, has limited the understanding of this crucial segment of higher education. While community colleges and other 2-year institutions often share a commitment to open access, comprehensiveness, and/or responsiveness to local needs, these unique American inventions are in fact a disparate group of institutions. The purpose of this report is to describe a classification system for 2-year institutions that can contribute to research and provide a framework for policy discussions.

Methodology

This study used the IPEDS database. The universe of institutions consists of 2,068 Title IV participating 2-year postsecondary institutions that met the data criteria for this study.[3] K-means cluster analysis—in combination with various other procedures—was the primary method used to classify these institutions.

---

[1]The Carnegie Foundation for the Advancement of Teaching. (2000). *The Carnegie Classification of Institutions of Higher Education.* Available: *http://www.carnegiefoundation.org/ Classification/CIHE2000/Tables.htm* [May 31, 2001].

[2] U.S. Department of Education, National Center for Education Statistics, 1995–96 National Postsecondary Student Aid Study (NPSAS:96), Data Analysis System.

[3] About 350 institutions were not included in this study because of inconsistencies in their data reporting in the different IPEDS surveys. For more information, see the Data Sources and Limitations section of the report.

Cluster analysis is a multivariate statistical procedure that attempts to mathematically form "clusters," or groups of relatively homogenous entities, based on measures of similarity with respect to specific variables, while maximizing the differences between groups. A focus group of experts in the field—researchers, association leaders, and policy analysts—along with preliminary data analysis, determined which variables were both policy relevant and appropriate to be included in the cluster analysis procedure. The results of the cluster analysis, in conjunction with subsequent analyses, revealed which variables created the most distinguishable categories of institutions.

The Classification System

The universe of institutions was first separated into three sectors by institutional control— public, private not-for-profit, and private for-profit. Further analyses were conducted within each sector, resulting in the seven-category classification system described below (figure A).
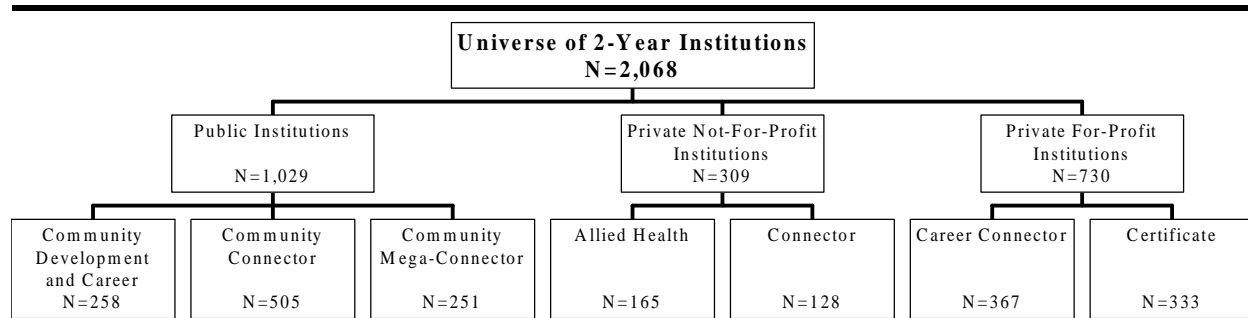
Public Institutions

Size of institutional enrollment is the most distinguishing characteristic of public 2-year institutions. Below is a brief description of the three categories within this sector:

Community Development and Career Institutions are institutions with an unduplicated headcount of less than 2,000 students. These institutions tend to confer awards and degrees primarily in job and career skills development, and to focus on overall workforce development for the communities they serve.

Community Connector Institutions are institutions with an unduplicated headcount of 2,000 to 9,999 students. These institutions tend to confer awards and degrees that target job and career skills development, and to offer academic programs with some component of general education that can facilitate transfer to 4-year institutions.

Figure A.—Classification system of 2-year institutions

NOTE: The sum of the number of institutions in each category does not add to the total number of institutions due to missing data in the variables chosen for categorization. In the universe of 2,068 institutions analyzed in this report, 61 institutions could not be placed in a final category: 15 public 2-year institutions, 16 private not-for-profit institutions, and 30 private for-profit institutions.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), 1997 Full Collection Year.

Community Mega-Connector Institutions are institutions with an unduplicated headcount of at least 10,000 students. These institutions tend to be in urban locations, to confer awards and degrees that target job and career skills development, and to offer academic programs with some component of general education that can facilitate transfer to 4-year institutions.

Private Not-For-Profit Institutions

The percentage of total awards granted that are in allied health programs is the distinguishing characteristic of private not-for-profit, 2-year institutions. Two categories were created within this sector:

Allied Health Institutions are institutions that grant 100 percent of their awards in allied health programs. These institutions tend to be small in enrollment and to have an exclusive focus on allied health training.

Connector Institutions are institutions that grant less than 100 percent of their awards in allied health programs. These institutions tend to confer awards and degrees that target job and career skills development, and to offer academic programs with some component of general education that can facilitate transfer to 4-year institutions.

Private For-Profit Institutions

A distinguishing characteristic of private for-profit 2-year institutions is the percentage of total awards granted that are certificates. Below is a description of the two categories within this sector:

Career Connector Institutions are institutions that grant less than 100 percent of their awards as certificates. They are degree-granting institutions—although many also offer certificates—that target job and career skills development. Many of these institutions offer academic programs with some component of general education that can facilitate transfer to 4-year institutions.

Certificate Institutions are institutions that grant 100 percent of their awards as certificates. These institutions provide specialized training, usually in a single job category or area.

Foreword

This report explores the development of a classification system for two-year institutions that can provide a framework for analysis and contribute to the discourse in public policy. The report discusses recent classification strategies put forth by researchers, and outlines a proposed classification system based upon nationally available data.

The report uses data collected under the Integrated Postsecondary Education Data System (IPEDS). IPEDS is a comprehensive data collection system that includes over 10,000 postsecondary accredited and non-accredited institutions including: baccalaureate or higher degree-granting institutions, 2-year award institutions, and less-than-2-year institutions. The IPEDS system is built around a series of interrelated surveys to collect institutional-level data on enrollment, program completion, faculty, staff, finances, and academic libraries.

In conducting the analysis for this report, the most recent year of final data available for each survey was used: 1997–98 Institutional Characteristics; 1997–98 Fall Enrollment; 1996–97 Completions; 1995–96 Finance; and 1997–98 Fall Staff. IPEDS defines a 2-year institution as a school with at least one program of at least 2 years but less-than-4 years in duration as their highest degree or award.

This study uses a methodology that includes both cluster analysis and descriptive statistics. "Cluster analysis" is the generic name for a variety of procedures that can be used to create a classification. These multivariate statistical procedures attempt to mathematically form "clusters" or groups of relatively homogenous entities based on measures of similarity and/or difference with respect to specific variables. Because of the large number of cases in this project—2,068 institutions—K-means cluster analysis was the appropriate method of cluster analysis. The K-means procedure begins by creating an aggregate mean for each case, and temporary estimates of the cluster means. Initial clusters are then formed by assigning each case to the cluster with the mean/center closest to its own, and then the cluster center is recalculated. An iterative process is used to find the final cluster centers, and at each step cases are grouped into the cluster with the closest center, and the cluster centers are recalculated. This process continues until no further changes are made in the centers or until a maximum number of iterations is reached. Additional information about the cluster analysis procedure and its use in the SPSS microcomputer application is contained in Appendix A.

Acknowledgments

THIS PAGE INTENTIONALLY LEFT BLANK

Table of Contents

Page

List of Tables

THIS PAGE INTENTIONALLY LEFT BLANK

List of Figures

THIS PAGE INTENTIONALLY LEFT BLANK

Introduction

A well-known classification systems for higher education in the United States was developed by the Carnegie Commission on Higher Education. Originally published in 1973—and refined and updated in 1976, 1987, 1994, and 2000—the Carnegie classification system divides colleges and universities into categories based on their degree-granting activity (doctoral, master's, baccalaureate, and associate's). In addition, the Carnegie classification system identifies myriad "specialized institutions," such as theological seminaries, schools of law, teachers colleges, and medical schools.

The evolution of this widely accepted classification system has made a substantial contribution to the development of literature and research in higher education. By identifying institutions that can be considered in the same "peer" group, a classification system enables researchers to conduct more rigorous analysis than otherwise would be possible and provides policymakers with reliable information. As a result, the Carnegie classification system has played a significant role in the formulation of higher education policy over the past three decades.

The Carnegie Foundation for the Advancement of Teaching has announced plans for a two-stage revision to the classification system. The preliminary revision was released in the fall of 2000 and the next revision is scheduled to coincide with the Foundation's centennial in 2005. That revision is intended to thoroughly overhaul the classification system by recognizing the many dimensions of institutional variation (McCormick 1999).

A limitation of the current Carnegie classification system is how 2-year institutions are categorized. While there are six distinct subcategories for 4-year institutions, the 2000 Carnegie classification system places the 1,669 institutions that offer associate of arts degrees or certificates and—with few exceptions—offer no baccalaureate degrees into a single grouping, Associate's Colleges. This single grouping of institutions accounts for more institutions than Doctoral/Research Universities, Master's Colleges and Universities, and Baccalaureate Colleges combined (1,478) (The Carnegie Foundation for the Advancement of Teaching 2000). Furthermore, the Carnegie classification includes only accredited and degree-granting institutions, excluding over 700 of the 2,427 2-year institutions with participation agreements for Title IV

student aid funds,[1] as found in the 1997–98 Institutional Characteristics survey of the Integrated Postsecondary Education Data System (IPEDS) database.

About half of all students in postsecondary education are enrolled in 2-year institutions.[2] Placing all 2-year institutions into one category, thereby masking their differences, has limited the understanding of this crucial segment of higher education. While community colleges and other 2-year institutions often share a commitment to open access, comprehensiveness, and/or responsiveness to local needs, these unique American inventions are in fact a disparate group of institutions. This diversity is reflected in a number of characteristics, including institutional control, enrollment, governance structure, geography (urban, rural, suburban), and types of degrees and certificates awarded. One example of diversity among 2-year institutions is the comparison of two community colleges. Northern Virginia Community College, which is located on the fringe of a large city, has an enrollment of over 35,000 students and a complex administrative structure. In comparison, Prince William Sound Community College is located in a small town in Alaska, has an enrollment of 750 students and a small administrative staff.[3]

The purpose of this report is to describe a classification system for 2-year institutions that can contribute to research and provide a framework for policy discussions. The report reviews recent classification strategies; discusses the data sources and limitations for the development of this classification system; explains the methodology, including a detailed description of cluster analysis; and, outlines a proposed classification system based upon nationally available data. In addition, an example of an actual institution in each category is provided to help understand the differences among them.

---

[1]These are institutions with participation agreements with the U.S. Department of Education for Title IV federal financial aid programs, based on the Postsecondary Education Participation System (PEPS) file.

[2]U.S. Department of Education, National Center for Education Statistics, 1995–96 National Postsecondary Student Aid Study (NPSAS:96), Data Analysis System.

[3]U.S. Department of Education, National Center for Education Statistics, 1997–1998 Integrated Postsecondary Education Data System (IPEDS), Fall Enrollment Survey.

Classification System Proposals

The *raison d'etre* of any classification system of postsecondary education institutions is to separate institutions into categories that are practical for the formulation of public policy and useful for educational researchers. In addition, a classification system must have a manageable number of categories and each institution must reside in one, and only one, category. The categories of a good classification system must differentiate the fundamental characteristics of the institutions, such as mission, curriculum, and/or student body. For example, separating 2-year institutions by the percentage of male faculty may be interesting and might even be supported by available data. However, the classification would be of little use to policy development because it is likely that the percentage of male faculty would have little relationship to mission, curriculum, and/or the student body of a 2-year institution. On the other hand, the size of student enrollment of a 2-year institution may be considered because it could discern or distinguish meaningful differences between institutions with regard to policy relevant characteristics. For example, as is shown later in this report, the size of student enrollment in public institutions is related to important characteristics such as the percentage of part-time faculty, the percentage of part-time students, and the percentage of awards granted that are certificates.

Several classification proposals found in the literature separate 2-year institutions in a variety of ways and in many categories. The following provides three examples of these schemas.

Classification by Control and Location

Using data from the U.S. Department of Commerce, the Bureau of the Census, and the IPEDS database, Stephen G. Katsinas proposed a classification scheme based upon institutional control (Katsinas 1996).[4] Three main categories were developed for 2-year institutions: 1) publicly controlled, 2) privately controlled, and 3) federally chartered and special use institutions. The publicly controlled institutions were further subdivided by location, into the categories of rural community colleges, suburban community colleges, and urban community colleges. The privately controlled institutions were separated into private not-for-profit, and proprietary.

Focusing upon the public sector of 2-year institutions, Katsinas offers several hypotheses that suggest major differences in several institutional characteristics of rural, suburban, and urban

---

[4]The universe of institutions used for this classification system is 2,501 public and private 2-year institutions.

community colleges. They include types of governance, physical plant, staffing, student race, student aid, remedial education, and workforce training and economic development. For instance, rural community colleges tend to be single-campus institutions that have centralized governance, while urban community colleges are generally multi-campus districts with highly decentralized governance. Suburban community colleges are a mix of single- and multi-campus institutions. Another example relates to administrative staffing. Rural community colleges often train their own financial affairs officers and lack sufficient funds for professional development. This contrasts with larger suburban community colleges that have the ability to hire professional staff and enjoy the benefits of professional training (Katsinas 1993).

Classification by Award Focus and Full-Time Enrollment

The National Center for Postsecondary Improvement (NCPI) proposed distinguishing 2-year institutions from one another by addressing the extent to which they focus on providing degrees and certificates, principally the associate's degree, versus providing a broad range of courses "to what is increasingly becoming a 'spot market' for educational services" (National Center for Postsecondary Improvement 1998). To capture the distinction among institutions that focus on degrees versus courses, NCPI used the following definitions for the categories.

1. *Degree Focus*: Full-time enrollment is at least 50 percent of an institution's total enrollment, and more than 15 percent of the students are awarded 2-year degrees and certificates each year.
2. *Mixed Focus*: Full-time enrollment is at 25 percent of total enrollment, and more than 10 percent (but less than 15 percent) of the students are awarded 2-year degrees and certificates each year.
3. *Course Focus*: All remaining institutions.[5]

Classification by Curricular Characteristics

In her doctoral dissertation, "A Curriculum-Based Classification System for Community Colleges," Gwyer Lenn Schuyler (1999) proposes categorizing public 2-year colleges based on course offerings, in particular, the percentage of institutional course offerings in the liberal arts.[6] Using the Spring college catalogs from 459 public community colleges, the study analyzed fifteen variables, including percentage of transferable course offerings, percentage of remedial course offerings, percentage of English course offerings, total revenues, and total enrollment.

---

[5]Using 1995 IPEDS data, 20 percent of the 1,094 institutions for which sufficient data were publicly available were classified as having a *Degree Focus*, with 40 percent each in the classifications of *Mixed Focus* and *Course Focus*.

[6]The universe of institutions used for this classification system is public institutions that are categorized by the 1994 Carnegie Classification as Associate of Arts colleges.

A total of thirteen classification models were tested; the model with the best fit was a system based on three categories of size and two categories of an English course proxy (percentage of English courses offered was used as a proxy for "Liberal Arts" and "Occupational" curricula). The categories included: Small Occupational, Small Liberal Arts, Medium Occupational, Medium Liberal Arts, Large Occupational, and Large Liberal Arts. However, as Schuyler notes, the ability to use this system is limited because it requires the collection of additional national level data from community colleges.

The best fitting system that uses data already collected for all institutions was a simple categorization based on enrollment size alone: large and small to medium. However, because of the strong correlation her research found between institutional size and curriculum (large institutions are very likely to offer liberal arts curriculums), Schuyler recommended that if this system were to be followed, it would be appropriate to name the two categories: 1) Large, Liberal Arts Colleges and 2) Small to Medium Colleges (with a range of curricular emphases).

THIS PAGE INTENTIONALLY LEFT BLANK

Data Sources and Limitations

The previous discussion of proposed classification systems draws attention to a major issue regarding the development of classification systems for 2-year institutions—data availability. In Schuyler's work, data were collected from catalogs from a sample of public community colleges; the challenge for this study is to use readily accessible national data that are useful for developing classifications. Establishing a meaningful classification system supportive of and necessary to policy development and research is limited to the extent to which appropriate data for the universe of 2-year institutions are collected at the national level.

It follows, then, that the data source for this project is the Integrated Postsecondary Education Data System (IPEDS). IPEDS is the core postsecondary education data collection program in the U.S. Department of Education's National Center for Education Statistics (NCES). It is a comprehensive data collection system developed to encompass *all* institutions and educational organizations whose primary purpose is to provide postsecondary education. Data are collected from over 10,000 postsecondary accredited and non-accredited institutions including: baccalaureate or higher degree-granting institutions, 2-year award institutions, and less-than-2-year institutions. The IPEDS system is built around a series of interrelated surveys to collect institutional-level data on enrollment, program completion, faculty, staff, finances, and academic libraries.

In conducting the analysis for this report, the most recent year of final data available for each survey was used: 1997–98 Institutional Characteristics; 1997–98 Fall Enrollment; 1996–97 Completions; 1995–96 Finance; and 1997–98 Fall Staff. The universe of institutions in the analysis is based on the 1997–98 Institutional Characteristics Survey and consists of the Title IV participating, 2-year institutions that either reported all data as an individual institution or reported consistently across the surveys as either a parent or child institution.[7] IPEDS defines a 2-year institution as a school with at least one program of at least 2 years but less-than-4 years in duration as their highest degree or award. In total, 2,068 of the 2,427 2-year institutions with Title IV participation agreements are included in the analysis.

---

[7]In IPEDS institutions can be identified as a parent or child institution, or an institution that submits data individually. A parent institution reports combined data, including another institution's data (the child institution) with their own. For example, an institution that is part of a state college system may report data individually or in combination with the other institutions in the system. Often institutions report their data inconsistently across surveys—for one survey they will report in combination with another institution (as a parent or child) while in a different survey they may report their data individually. This makes analysis with these institutions difficult, therefore those institutions that reported inconsistently were not included in the universe for this analysis. For more information, see Appendix A of the report.

Data Limitations

Due to limitations in the IPEDS data, especially with respect to information on 2-year institutions, there are several aspects of these institutions that could not be examined. IPEDS surveys collect more data from degree-granting institutions than from 2-year institutions that only offer certificates. For example, the age of enrolled students and many financial data elements are only collected from degree-granting institutions. Unfortunately, this restricted the analysis that could be performed on non-degree granting institutions—almost 30 percent of the universe.[8]

Aspects of these institutions that could help to distinguish mission are not included in IPEDS data collection. Lists of programs offered are not available[9] and non-credit course work is not captured either, even though non-credit activity is often a large part of the curriculum and student enrollment at many 2-year institutions. In addition, many 2-year institutions are part of a multi-campus system of institutions while others are single campus institutions, which is difficult to determine in IPEDS since institutional governance is not a variable in the dataset. These aspects of 2-year institutions could help to distinguish these schools from one another, however these data are not collected on the national level.

---

[8]Though non-degree granting 2-year institutions account for almost 30 percent of the universe of 2-year institutions, it is important to note that only approximately 4 percent of the enrollment in 2-year institutions is in non-degree granting schools.

[9]NCES collected data for several survey cycles on Fall Enrollment in Occupationally Specific Programs (EP Survey), which could help capture curriculum at 2-year institutions, but the data were considered incomplete and therefore not released. These data have not been collected from institutions since 1997.

Cluster Analysis

There are many possible methods of analysis to arrive at a classification scheme; this study uses a combination of cluster analysis and descriptive statistics. Clustering methods may not be as clearly understood in comparison to other, more common statistical procedures. Because the literature often reflects contradictory methods and preferred approaches, it is useful to discuss general cluster analysis procedures before focusing on the specifics used in this project's analysis.

"Cluster analysis" is the generic name for a variety of procedures that can be used to create a classification. These multivariate statistical procedures attempt to mathematically form "clusters" or groups of relatively homogenous entities based on measures of similarity and/or difference with respect to specific variables. Though many methods exist, hierarchical and K-means (iterative) cluster analysis are the most widely used. The hierarchical clustering method, however, is not as appropriate for a large number of cases, as the results become unwieldy. Therefore, because of the large number of cases in this project—2,068 institutions—K-means was a more appropriate method of cluster analysis.

The K-means procedure begins by creating an aggregate mean—combining all variables included in the analysis—for each case (in this project, for each institution) and then temporary estimates of the cluster means.[10] Initial clusters are then formed by assigning each case to the cluster with the mean/center closest to its own, and then the cluster center is recalculated. An iterative process is used to find the final cluster centers, and at each step cases are grouped into the cluster with the closest center, and the cluster centers are recalculated. This process continues until no further changes are made in the centers or until a maximum number of iterations is reached.

K-means cluster analysis requires the specification of the number of clusters to be formed. Often the "natural" or optimal number of clusters is not known, therefore methods have been developed to help determine this number. The most common procedure is to run a subset of cases in hierarchical cluster analysis and look for "jumps" in the fusion coefficient—the numerical value at which various cases merge to form a cluster. A "jump" in the fusion coefficient suggests

---

[10]The values of the first k cases in the data file are used as temporary estimates of the k cluster means, where k is the number of clusters that are to be formed. The number of clusters to be formed is specified by the user (SPSS Inc. 1999).

that two relatively dissimilar clusters have been merged; thus the number of clusters prior to the merger is the most probable solution. Another appropriate strategy is to try several different analyses (for example, requesting three, four, and five clusters) in a search for the most appropriate solution. Either way a judgment about the number of clusters must be made; unfortunately, there is no single test that reveals the exact number of clusters that should be generated (Aldenderfer and Blashfield 1984). For the purposes of this report, both methods were used to help guide the K-means cluster analysis. Hierarchical analysis was used to find an appropriate range for the number of clusters, and those cluster numbers within the range were all tried in the analysis to determine which was the most appropriate.

In SPSS, a K-means analysis result produces the distance each case is from its cluster center as well as an ANOVA table. The size of the "F" statistic—the ratio of the between-cluster mean square and the within-cluster mean square—in the K-means one-way ANOVA is used for identifying variables that drive the clustering and those that differ little across clusters. In cluster analysis, the "F" statistic is not used to test significant differences between groups—the clusters are formed to characterize difference. In short, the "F" statistics provide information about each variable's contribution to the separation of the groups; once the driving variables have been identified, they can be used to create meaningful categories.

The choice of variables to be included in the cluster analysis is one of the most critical steps in the process and probably has "the greatest influence on the ultimate results of a cluster analysis" (Anderberg 1973). Because the analysis uses an aggregate mean, each variable that is included in the analysis affects the clustering results—this is one of the reasons why the choice of variables is crucial. Ideally, variables should be chosen within the context of a theory used to support the classification and serve as the basis for the choice of variables to be used. As discussed in the methodology section, this study used a combination of a focus group and preliminary analysis in order to choose appropriate variables.

An important limitation of K-means cluster analysis is that a categorical variable should not be run in this form of cluster analysis. K-means analysis creates clusters based on the means of variables, but finding the mean of a categorical variable is inappropriate.[11] In addition, careful attention must be paid to correlations between the chosen variables, since the use of highly correlated variables is essentially an implicit "weighting" of these variables; a phenomenon which overstates the value of the variable or construct underlying the correlated set (Aldenderfer and Blashfield 1984). If three highly correlated variables are included, the effect is similar to using

---

[11]Ordinal variables are not ideal for K-means cluster analysis either, however, they are less problematic than categorical variables because they do have order, giving the values more meaning. Urbanicity/location is an ordinal variable, and was included in the cluster analysis.

only one variable that has a weight three times greater than any other variable. In general, highly correlated variables should not be simultaneously run in cluster analysis.

It is also important to note that during the cluster analysis process, certain clusters may be so "natural and self-evident as to . . . be revealed by almost any method" (Anderberg 1973). When it is the case that certain variables or clusters present themselves as obvious candidates for categorization, it is appropriate to remove either the variable or the group of cases and continue the analysis with the remaining sub-set(s) of data. In this study, the preliminary analysis—discussed in the next section—is used carefully to study the descriptive statistics of the variables considered for analysis and as a result, to discover any "natural" clusters.

THIS PAGE INTENTIONALLY LEFT BLANK

Methodology

For this study, cluster analysis was supplemented with various procedures in order to find the "best" variables and categorization. These procedures are described below.

The Literature Review and Focus Group Meeting

As discussed in the previous section, one of the most critical steps in cluster analysis procedure is choosing which variables to include in the analysis. A literature review of past research addressing the classification of 2-year institutions provided the initial direction for considering variables. To add to the information gathered from past research, a focus group meeting was held. A group of seven experts—researchers, association leaders, and policy analysts—familiar with 2-year institutions was convened in December 1999; they helped to guide the selection of appropriate variables and their definitions. The focus group provided a starting point for the analysis by offering feedback on potential variables for classification based on their availability, usefulness for public policy purposes, and relevance to the 2-year postsecondary education community.

The Preliminary Analysis

Twenty potential variables were selected by the focus group. Those variables that were available through IPEDS were then further investigated and refined based on examinations of their descriptive statistics through frequency distributions and histograms. First and foremost, variables had to have sufficient data. When approximately 30 percent of the institutions in the universe lacked data for the variable, the data were considered insufficient and were not considered for further analysis. Two other major criteria guided the remaining investigation of the variables: *range and variance*, and *replicability*.[12] Each variable had to show a wide range and variance among the institutions so that there was the potential to create distinguishable categories. In addition, the variables and classification system as a whole had to be replicable so that the classification system designed would continue to be useful in the future and could be refined and

---

[12]Due to changes made in the IPEDS Finance Survey, the 1995–96 data used in this study may not be comparable to future reported data. Still, because of the potential importance of finance variables, they were included in the analysis.

updated when necessary. Variables that did not meet these requirements were eliminated from further analysis.

The Chosen Variables

The following is a list of the variables that remained either after the preliminary analysis on the universe of institutions or the preliminary analyses that were later conducted on each sector of 2-year institutions (definitions of the variables are presented in table 1). It is important to note that throughout the process, variables were re-analyzed for relevance within each subgroup formed (figure 1).

- Urbanicity/location
- 12-month unduplicated headcount/enrollment
- Percentage of:
  - full-time, first-time degree-seeking students
  - part-time students
  - minority students
  - non-traditional aged students
  - part-time faculty
  - awards granted as certificates
  - awards granted in allied health programs
  - awards granted in occupationally specific programs[13]
  - educational & general (E&G) expenditures for instruction
  - E&G expenditures for scholarships and fellowships
  - revenues from state and local support
- Institutional control

The variable "institutional control" was considered for classification based on the frequency distribution, range and variance, and cross-tabulations with other variables. Institutional control was used as a blocking variable to define 3 groups—public, private not-for-profit, and private for-profit—that were then analyzed independently.

---

[13]Although it is important to note that many students at 2-year colleges do not receive a degree, but instead attend for short-term training and vocational purposes, or to transfer to another institution—program completions/awards granted are still reflective of the curriculum of an institution.

Table 1.—Definitions of variables that remained after at least one of the preliminary analyses and subsequently were included in one or more stages of cluster analysis

Variable	Definition
Urbanicity/location	Indicates the degree of urbanization of the institution's location.
	1. Large City—A central city of a Consolidated Metropolitan Statistical Area (CMSA) or Metropolitan Statistical Area (MSA), with the city having a population greater than or equal to 250,000.
	2. Mid-size City—A central city of a CMSA or MSA, with the city having a population less than 250,000.
	3. Urban Fringe of Large City—Any incorporated place, Census Designated Place (CDP), or non-place territory within a CMSA or MSA of a Large City and defined as urban by the Census Bureau.
	4. Urban Fringe of Mid-size City—Any incorporated place, CDP, or non-place territory within a CMSA or MSA or a Mid-size City and defined as urban by the Census Bureau.
	5. Large Town—An incorporated place or CDP with a population greater than or equal to 25,000 and located outside a CMSA or MSA.
	6. Small Town—An incorporated place or CDP with a population less than 25,000 and greater than or equal to 2,500 and located outside a CMSA or MSA.
	7. Rural—Any incorporated place, CDP, or non-place territory designated as rural by the Census Bureau.
12-month unduplicated headcount	Indicates the unduplicated headcount of undergraduates enrolled during the 12-month period.
Percentage of full-time, first-time degree-seeking students	Indicates full-time, first-time degree seeking undergraduate enrollment (12-Month unduplicated headcount) as a percentage of total undergraduate enrollment (12-month unduplicated headcount).
Percentage of part-time students	Indicates part-time undergraduate enrollment (fall enrollment) as a percentage of the total fall enrollment.
Percentage of minority students	Indicates minority student enrollment (fall enrollment) as a percentage of the total fall enrollment. "Minority" includes students who are Black, non-Hispanic, American Indian or Alaskan Native, Asian or Pacific Islander, or Hispanic.
Percentage of non-traditional aged students	Indicates non-traditional aged student enrollment (fall enrollment) as a percentage of total fall enrollment. "Non-traditional aged" includes students age 25 and older. This data is only collected for degree-granting institutions.
Percentage of part-time faculty	Indicates part-time faculty as a percentage of total faculty on staff.
Percentage of awards granted as certificates	Indicates number of certificates granted as a percentage of total awards granted.
Percentage of awards granted in allied health programs	Indicates awards, certificates or degrees granted in allied health programs as a percentage of the total awards granted. "Allied health" includes all programs identified by NCES' Classification of Instructional Programs (CIP) as "Health Professions and Related Sciences" (CIP code of 51).

Table 1.—Definitions of variables that remained after at least one of the preliminary analyses and subsequently were included in one or more stages of cluster analysis—Continued

Variable	Definition
Percentage of awards granted in occupationally specific programs	Indicates awards, certificates, or degrees, granted in occupationally specific programs as a percentage of total number of awards granted. "Occupationally specific" includes all programs identified by NCES' Classification of Instructional Programs (CIP) as "instructional programs whose expressed intent is to impart work-related knowledge and skills at the secondary and postsecondary levels."
Percentage of education and general (E&G) expenditures for instruction	Indicates expenditure on instruction as a percentage of the total current fund E&G expenditures.
Percentage of E&G expenditures for scholarships and fellowships	Indicates expenditures on scholarships and fellowships as a percentage of the total current fund E&G expenditures.
Percentage of revenues from state and local support	Indicates revenue that is from both state and local appropriations, grants, and contracts as a percentage of the total current funds revenue.
Institutional control	Indicates the control of the institution—public, private not-for profit, and private for-profit.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), Full 1997 Collection Year.

Cluster Analysis With 2-Year Institutions

Correlated variables were determined from both correlation matrices and hierarchical cluster analyses of the variables themselves.[14] Highly correlated variables were run in separate cluster analyses in order to avoid overstating their value in the analysis, as discussed in the previous cluster analysis section.

Because the K-means method of clustering demands that the number of clusters in the analysis be determined in advance and an optimal number of clusters was not obvious, two methods were used to help determine an appropriate number. Hierarchical cluster analysis was performed on a sample of the institutions in the universe and the fusion coefficients were examined (refer to the cluster analysis section and Appendix A for more detailed descriptions of this

---

[14]Hierarchical cluster analysis can be used to cluster either cases or variables. In this case, it was used to cluster variables in order to help identify related groups of variables. The hierarchical method begins by finding the closest pair of variables (or cases) according to the determined distance measure and combines them to form a cluster. The algorithm continues joining pairs of variables (or cases), pairs of clusters, or a variable (or case) with a cluster, until all the data are in one cluster. The method is *hierarchical* because once two variables (or cases) are joined, they remain so until the final step—a cluster formed in a later stage of the analysis contains clusters from earlier stages and so on. Variables that are joined at the earliest stages of the analysis are considered to have the strongest relationships (SPSS Inc. 1999).

Figure 1.—Flow chart of the derived variables for categorization of 2-year institutions

Variables from the literature review and focus group meeting

- Urbanicity
- 12-month unduplicated headcount
- Percentage of full-time, first-time degree-seeking students
- Percentage of part-time students
- Percentage of minority students
- Percentage of non-traditional aged students
- Percentage of female students
- Percentage of part-time faculty
- Percentage of awards granted as certificates
- Percentage of awards granted in allied health programs
- Percentage of awards granted in occupationally specific programs
- Percentage of E&G expenditures for instruction
- Percentage of E&G expenditures for scholarships and fellowships
- Percentage of E&G expenditures for public service
- Percentage of revenues from state and local support
- Percentage of revenues from federal grants and contracts
- Institutional control
- Institutional governance
- Institutional affiliation
- Degree-granting status

Variables remaining after the preliminary analysis on the universe of institutions

- Urbanicity
- 12-month unduplicated headcount
- Percentage of full-time, first-time degree-seeking students
- Percentage of part-time students
- Percentage of minority students
- Percentage of non-traditional aged students
- Percentage of part-time faculty
- Percentage of awards granted in occupationally specific programs
- Percentage of awards granted as certificates
- Percentage of E&G expenditures for instruction
- Percentage of revenues from state and local support
- Institutional control *(not included in the cluster analysis)*

Top 3 variables from the initial cluster analysis on the universe of institutions

- Percentage of revenues from state and local support
- 12-month unduplicated headcount
- Percentage of part-time students

These variables were considered a proxy for institutional control.

Variables after the preliminary analysis on public institutions

Variables after the preliminary analysis on private not-for-profit institutions

Variables after the preliminary analysis on private for-profit institutions

Figure 1.—Flow chart of the derived variables for categorization of 2-year institutions—Continued

Variables after the preliminary analysis on public institutions	*Variables after the preliminary analysis on private not-for-profit institutions*	*Variables after the preliminary analysis on private for-profit institutions*
• Urbanicity • 12-month unduplicated headcount • Percentage of full-time, first-time degree-seeking students • Percentage of part-time students • Percentage of minority students • Percentage of non-traditional aged students • Percentage of part-time faculty • Percentage of awards granted as certificates • Percentage of awards granted in allied health programs • Percentage of awards granted in occupationally specific programs • Percentage of revenues from state and local support	• Urbanicity • Percentage of full-time, first-time degree-seeking students • Percentage of part-time students • Percentage of minority students • Percentage of non-traditional aged students • Percentage of part-time faculty • Percentage of awards granted as certificates • Percentage of awards granted in allied health programs • Percentage of awards granted in occupationally specific programs • Percentage of E&G expenditures for instruction	• Percentage of full-time, first time degree-seeking students • Percentage of minority students • Percentage of part-time faculty • Percentage of awards granted as certificates • Percentage of E&G expenditures for instruction • Percentage of E&G expenditures for scholarships and fellowships

Top variables from cluster analysis on public institutions	Top variables from descriptive analysis on private not-for-profit institutions	Top variables from cluster analysis on private for-profit institutions
• Urbanicity • 12-month unduplicated headcount • Percentage of awards granted as certificates • Percentage of part-time students	• Percentage of awards granted in allied health programs	• Percentage of awards granted as certificates • Percentage of full-time, first-time degree-seeking students • Percentage of minority students • Percentage of part-time faculty

Final variable from post-analysis on public institutions	Final variable from analysis on private not-for-profit institutions	Final variable from post-analysis on private for-profit institutions
• 12-month unduplicated headcount	• Percentage of awards granted in allied health programs	• Percentage of awards granted as certificates

method).[15] Using this information, a series of different analyses were run varying the number of clusters while also changing the combination of variables to avoid running highly correlated variables in the same analysis. This led to the examination of many different variations of the analyses, making it possible to determine if certain variables were consistently assigned the highest "F" ratios and were therefore the "driving" variables for separating the institutions from one another.

The Post-Analysis

Once the cluster analysis was completed, "driver" variables were considered both on their own and in combination with other variables through the use of scatterplots, crosstabulations, and frequency distributions. The characteristics (mean/median of all the variables) of the different groups created by these "driver" variables also were examined to determine which categories were most different from each other. Those variables that created the most distinct categories were chosen to guide the classification system.

It is important to emphasize that the seven categories presented in this classification system are not the actual clusters produced by the cluster analysis. The clusters that are produced by K-means cluster analysis are based on aggregate means. The "value" of the means that result from the analysis are strictly operational, and have no application for institutions and policymakers. Rather, the analysis identified the "driver" variables—those with the highest "F" statistics—that were then further examined through the post-analysis. From the post-analysis, the "best" variable(s) was determined and then used to separate the institutions into the different categories of the classification system. Consistent with standard cluster analysis procedure, once the variable(s) for classification were identified, the entire cluster analysis process was then conducted within the subgroups formed (figure 1).

---

[15]From the fusion coefficients it appeared that between six and eight clusters were an appropriate number for the initial cluster analysis. This procedure was also used for the sub-categories of public and private-for-profit institutions in order to determine an appropriate number of categories for those sectors. However, it was not used for the private, not-for profit institutions because the sector includes such a small group of institutions, that two categories were appropriate.

THIS PAGE INTENTIONALLY LEFT BLANK

The Classification System

The following narrative describes the classification system for 2-year institutions and out-lines the rationale underpinning its development. A description of the categories of 2-year institutions is provided along with the characteristics of each. For each category, an example of an actual institution is presented.

Institutional Control of 2-Year Institutions is a Major Factor

The results of the analysis across the universe of 2-year institutions showed that institutional control is a distinguishing characteristic of these schools. The cluster analysis pointed towards three highly correlated variables: the percentage of revenues from state and local appropriations, grants, and contracts; institutional enrollment; and the percentage of students enrolled part-time. All three of these variables are highly correlated with institutional control (and with each other)—a variable that could not be included in the cluster analysis because of its categorical nature. This strongly indicated that the three variables were serving as a proxy for control (table 2). Based on the crosstabulations with other variables as well as the frequency distributions, it was evident that control was a strong variable for categorizing these institutions. The three categories are: public institutions; private not-for-profit institutions; and private for-profit institutions (figure 2).

Table 2.—Pearson product-moment correlations among selected variables

	Institutional control	12-month unduplicated headcount	Percentage of part-time students	Percentage of revenues from state and local support
Institutional control	1.00	-0.43	-0.68	-0.85
12-month unduplicated headcount	-0.43	1.00	0.50	0.42
Percentage of part-time students	-0.68	0.50	1.00	0.70
Percentage of revenues from state and local support	-0.85	0.42	0.70	1.00

NOTE: Correlations are significant at the .01 level (2-tailed).

SOURCE: U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), Full 1997 Collection Year.

Figure 2.—Frequency distribution of 2-year institutions by institutional control
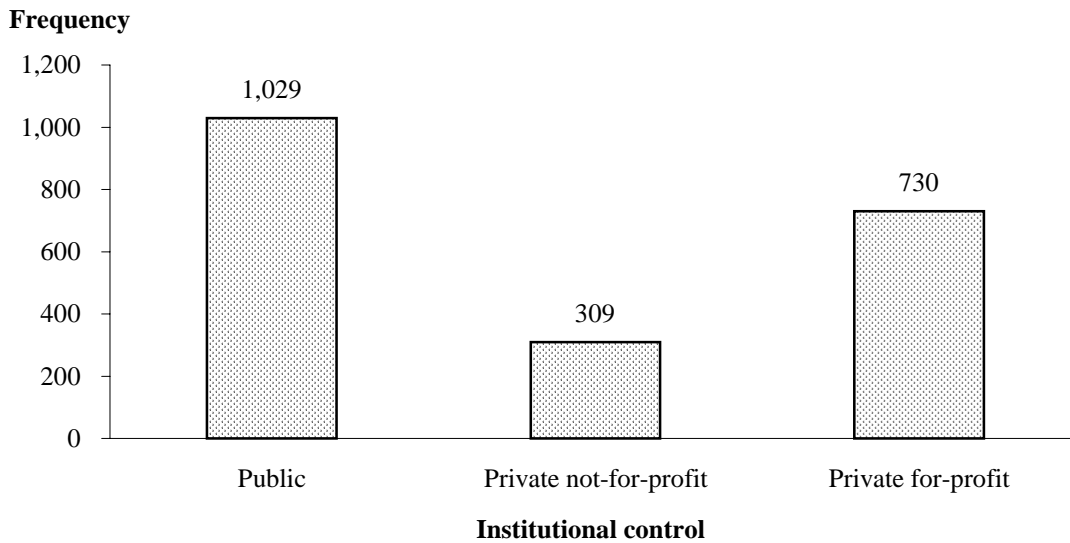
Frequency

SOURCE: U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), 1997 Full Collection Year.

In the remainder of the report, generalizations are made about the institutions that fall into each category. These generalizations are derived from two sources: (1) the descriptive statistics of variables from the IPEDS data (table 3) and (2) from a review of the actual institutions in each category. It is important to note that these generalizations may not be true for all institutions in that category but are intended to help better understand how the categories differ.

Public Institutions (N=1,029)

Public institutions have a higher median enrollment (4,318 students) than the other two sectors of institutions in addition to having a higher median percentage of part-time students (58 percent). These institutions have a lower median percentage of full-time, first-time degree-seeking students enrolled (12 percent) and a lower median percentage of awards granted that are certificates (30 percent)—the lowest of all the sectors. Institutions in the public category also have a higher median percentage of part-time faculty or staff (61 percent) than both categories of private institutions, and not surprisingly, have a much higher median percentage of current fund revenues from state and local support (56 percent). Public 2-year institutions are likely to be located around an urban area (median=3/urban fringe of a large city).

Table 3.—Characteristics of the institutional control categories for 2-year institutions, according to selected variables

	Public N=1,029		Private not-for-profit N=309		Private for-profit N=730	
	Mean	Median	Mean	Median	Mean	Median
Urbanicity	—	3.0	—	2.0	—	2.0
12-month unduplicated headcount	7,750.4	4,318.0	326.5	138.0	413.3	254.0
Percentage of full-time, first-time degree-seeking students	18.5	12.2	47.2	41.3	62.1	62.8
Percentage of part-time students	54.8	58.4	18.1	4.2	13.2	0.0
Percentage of minority students	26.3	20.0	21.1	11.5	29.7	22.1
Percentage of non-traditional aged students	45.7	46.8	39.9	39.6	42.4	42.0
Percentage of part-time faculty	49.9	61.2	31.8	24.2	36.7	33.3
Percentage of awards granted as certificates	39.3	29.9	62.6	100.0	65.6	95.9
Percentage of awards granted in allied health programs	25.1	22.9	61.0	100.0	12.9	0.0
Percentage of occupationally specific awards	71.1	72.5	84.9	100.0	92.9	100.0
Percentage of E&G expenditures for instruction	45.7	44.1	48.5	42.4	31.6	27.2
Percentage of revenues from state and local support	56.9	56.5	5.0	2.6	5.1	0.0

— Not applicable.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), Full 1997 Collection Year.

Private Not-For-Profit Institutions (N=309)

The private not-for-profit institutions have the smallest median enrollment of the three sectors (138 students), a low median percentage of part-time students (4 percent), and tend to be located in more urban locations (median=2/mid-size city). This group of institutions has a median of 100 percent for the percentage of awards granted that are certificates, the percentage of awards granted that are in occupationally specific programs, and the percentage of awards granted in allied health fields.

Private For-Profit Institutions (N=730)

Like the private not-for-profit institutions, the private for-profit group of institutions tend to be located in more urban locations (median=2/mid-size city). These institutions fall in the middle of the three sectors with respect to median enrollment (254 students) but have the lowest median percentage of part-time students (0 percent). On the other hand, they tend to have a high median percentage of full-time, first-time degree-seeking students enrolled (63 percent)—the highest of the three sectors. Like the not-for-profits, these institutions have a median of 100 percent awards granted in occupationally specific programs, and they have a high median percentage of awards granted that are certificates (96 percent). However, of the three sectors, the for-profit institutions

have an extremely low median percentage of awards granted in allied health programs (0 percent), and a relatively low percentage of E&G expenditures for instruction (27 percent).

The Remaining Analysis

All further analyses, including both cluster analysis and the examination of the descriptive statistics, were conducted separately on each of the subgroups described above. Each variable was examined again, based on *range and variance*, to determine if it was appropriate for the particular subgroup of institutions that was being investigated (figure 1). Correlated variables in each subgroup were also determined so as to avoid running them in the cluster analyses simultaneously. The resulting classification system consists of the following categories: three categories for public institutions; two categories for private not-for-profit institutions; and two categories for private for-profit institutions, labeled as follows:

Public Institutions

- Community Development and Career Institutions
- Community Connector Institutions
- Community Mega-Connector Institutions

Private Not-For-Profit Institutions

- Allied Health Institutions
- Connector Institutions

Private For-Profit Institutions

- Career Connector Institutions
- Certificate Institutions

Public 2-Year Institutions

Institutional enrollment is the most distinguishing characteristic of public 2-year institutions and is clearly the driving variable for differentiating public institutions from one another. The cluster analysis, in conjunction with post-analyses, indicated that institutional enrollment on its own is the "best" at creating categories with different characteristics—other variables did not appear to add to the differentiation of these institutions. After examining many different options for separating the institutions by enrollment, using the 25th and 75th percentiles seemed a rea-

sonable option as it clearly created categories that differed from one another (table 4; figure 3). For ease of analysis and policy discussion, these percentiles were rounded to 1,999 students (less than 2,000 students) and 9,999 students (less than 10,000 students).

Once the categories were determined, they were named by examining the characteristics of the institutions in each category.[16] As noted above, the names for all seven categories were developed by analysis of the IPEDS variables and a review of the actual institutions that reside in the category. Three distinguishable categories were created for public 2-year institutions: *Community Development and Career Institutions*; *Community Connector Institutions*; and *Community Mega-Connector Institutions* (table 5).

Each category is described below with respect to the variables in the analysis, followed by a more interpretive description of the category (based both on the data and on an examination of the literature and the institutions within each category). Finally an example of an institution that falls near the median of the category is presented (figure 4).
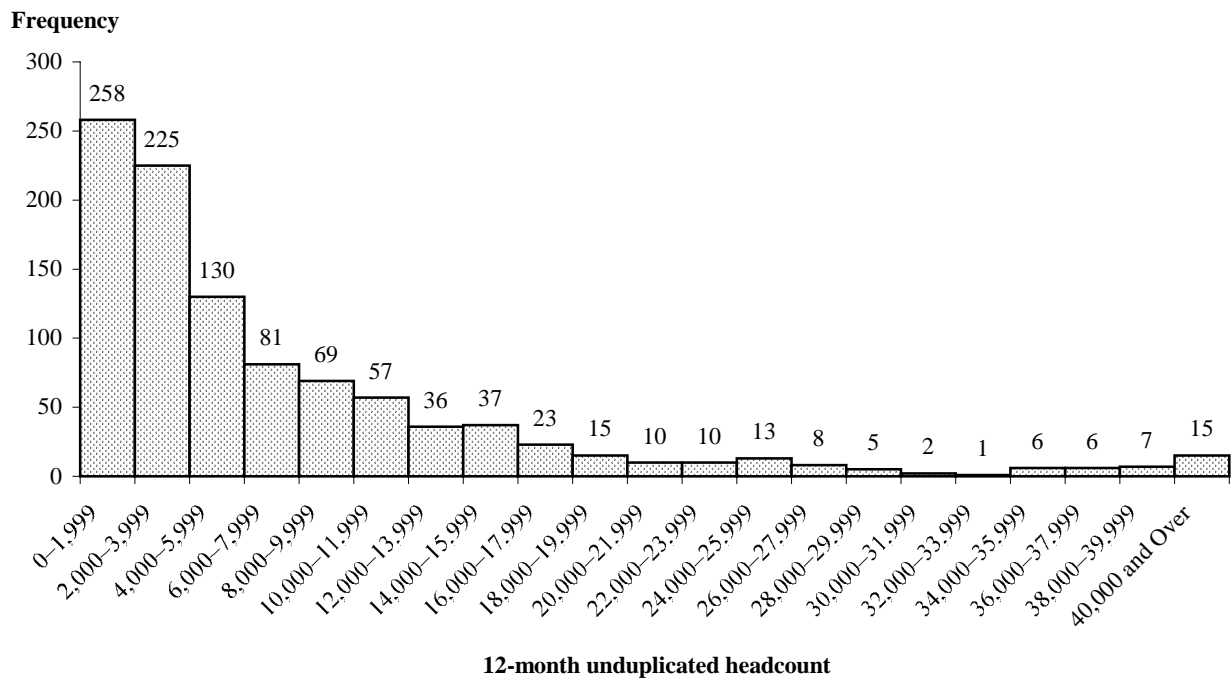
Table 4.—Descriptive statistics of the 12-month unduplicated headcount at public, 2-year institutions

N	
Valid	1,014.0
Missing	15.0
Mean	7,750.4
Median	4,318.0
Standard Deviation	10,052.4
Variance	101,000,000.0
Range	126,021.0
Minimum	7.0
Maximum	126,028.0
Percentiles	
25%	1,919.5
50%	4,318.0
75%	9,909.0

SOURCE: U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), Full 1997 Collection Year.

---

[16]When naming the categories that were formed, an emphasis was placed on creating names that were not value-laden. One of the goals was to design a classification system for 2-year colleges that was non-hierarchical, so that one category would not be seen as more desirable than another.

Figure 3.—Frequency distribution of public, 2-year institutions by the 12-month unduplicated headcount

Frequency

300 –

258
225
130
81
69
57
36 37
23
15
10 10 13
8
5
2
1
6 6
7
15

0–1,999
2,000–3,999
4,000–5,999
6,000–7,999
8,000–9,999
10,000–11,999
12,000–13,999
14,000–15,999
16,000–17,999
18,000–19,999
20,000–21,999
22,000–23,999
24,000–25,999
26,000–27,999
28,000–29,999
30,000–31,999
32,000–33,999
34,000–35,999
36,000–37,999
38,000–39,999
40,000 and Over

12-month unduplicated headcount

SOURCE: U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), 1997 Full Collection Year.

Community Development and Career Institutions (N=258)

The *Community Development and Career Institutions* category includes institutions with less than 2,000 students enrolled. These institutions tend to be in less urban locations than the larger *Community Connector* and *Community Mega-Connector Institutions*—the median location for institutions in this category is a "small town" (median=6/small town). *Community Development and Career Institutions* have a lower median percentage of part-time students enrolled (42 percent) and a higher median percentage of full-time, first-time degree-seeking students enrolled (25 percent) than the other categories of public institutions. There is a tendency for a higher percentage of awards granted at these institutions to be certificates (median=82 percent) and a higher percentage of awards granted in occupationally specific programs (median=100 percent) than at institutions in the other two categories. Finally, *Community Development and Career Institutions* have a smaller median percentage of part-time faculty (23 percent) than do the *Community Connector* and *Community Mega-Connector Institutions*.

Table 5.—Characteristics of the categories for public, 2-year institutions, according to selected variables

	Mean	Median
Community Development and Career (N=258)		
Urbanicity	—	6.0
12-month unduplicated headcount	940.2	938.5
Percentage of full-time, first-time degree-seeking students	31.4	24.5
Percentage of part-time students	39.2	42.3
Percentage of minority students	29.4	22.6
Percentage of non-traditional aged students	42.9	45.6
Percentage of part-time faculty	31.0	22.7
Percentage of awards granted as certificates	66.3	82.1
Percentage of awards granted in occupationally specific programs	83.9	100.0
Percentage of E&G expenditures for instruction	49.8	47.0
Community Connector (N=505)		
Urbanicity	—	4.0
12-month unduplicated headcount	4,947.0	4,372.0
Percentage of full-time, first-time degree-seeking students	15.6	12.1
Percentage of part-time students	56.8	58.6
Percentage of minority students	22.2	16.2
Percentage of non-traditional aged students	44.9	45.4
Percentage of part-time faculty	52.7	61.8
Percentage of awards granted as certificates	32.2	25.8
Percentage of awards granted in occupationally specific programs	69.5	71.3
Percentage of E&G expenditures for instruction	44.2	43.0
Community Mega-Connector (N=251)		
Urbanicity	—	2.0
12-month unduplicated headcount	20,390.7	15,533.0
Percentage of full-time, first-time degree-seeking students	9.0	7.1
Percentage of part-time students	66.2	68.6
Percentage of minority students	31.5	25.8
Percentage of non-traditional aged students	49.0	49.0
Percentage of part-time faculty	62.4	68.7
Percentage of awards granted as certificates	26.8	23.1
Percentage of awards granted in occupationally specific programs	61.6	60.4
Percentage of E&G expenditures for instruction	45.0	44.4

— Not applicable.

NOTE: The sum of the number of institutions in each category does not add to the total number of public institutions due to missing data in the variables used for categorization.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), Full 1997 Collection Year.

Figure 4.—Characteristics of the public, 2-year institution examples, according to selected variables
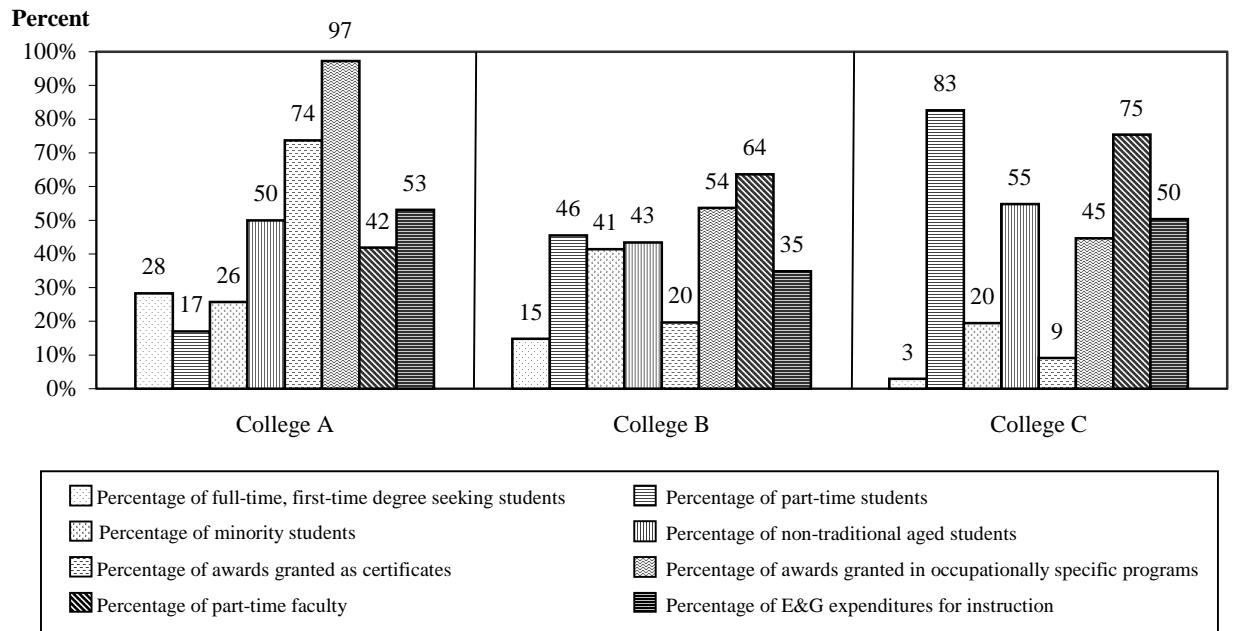
SOURCE: U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), 1997 Full Collection Year.

More interpretively, *Community Development and Career Institutions* tend to be schools that confer awards and degrees primarily in job and career skills development for students, and to focus on overall workforce development for the communities they serve. Most of these schools are located in rural areas, often not served by proprietary schools. Therefore, these institutions frequently offer certificate programs usually found in private for-profit schools that tend to be located in more urban areas. In addition to their education services, these institutions may serve a cultural role in their communities as well. Schuyler's research, which found a strong correlation between enrollment and curriculum, supports the idea that these smaller public institutions are likely to have more occupational curricula than are the larger institutions (Schuyler 1999).

Institution A[17] is an example of an institution that falls into the *Community Development and Career Institution* category. This institution has an enrollment of 1,044 students and is located in a small town. The mission of the institution is to provide occupational education for those who wish to prepare for or retrain in a career field. Though this institution does offer associate's degrees in some technical programs, the majority of awards granted are certificates in

---

[17]Examples given are actual institutions that have been assigned aliases for the purpose of anonymity. Descriptions of the example institutions are based on a combination of IPEDS data and published documents about the institutions.

programs such as auto body repair, carpentry, cosmetology, and interior design. One of the stated goals of the institution is to promote economic development in the area by providing training and instruction in courses designed to meet the specific needs of business and industry.

Community Connector Institutions (N=505)

The *Community Connector Institutions* consist of schools with enrollments of 2,000 to 9,999 students; with 505 institutions, this is the largest category in the classification system.[18] These institutions are generally not found in as urban locations as the *Community Mega-Connector Institutions*, but tend to be in more urban locations than the *Community Development and Career Institutions*—the median location (4) is defined as an "urban fringe of a mid-size city. *Community Connectors* fall in the middle of the three categories for median percentage of part-time students (59 percent) and for median percentage of full-time, first-time degree-seeking students (12 percent). A lower median percentage of minority students are enrolled (16 percent) at these institutions compared to both other categories of public institutions. Compared to the *Community Development and Career Institutions*, these schools have a low median percentage of their awards granted as certificates (26 percent)—similar to the median percentage of the *Community Mega-Connectors*. These institutions fall in the middle of the three categories for median percentage of their awards granted in occupationally specific programs (71 percent). Finally, *Community Connectors* have a higher median percentage of part-time faculty (62 percent) on staff as compared to the *Community Development and Career Institutions*.

When looking at the institutions in this category, it appears that they tend to confer awards and degrees that target job and career skills development, and to offer academic programs with some component of general education that can facilitate transfer to 4-year institutions. These institutions tend to enroll a large number of part-time students and often have a high percentage of part-time faculty on staff. In addition, they offer a wide variety of programs and may be important feeder schools for colleges and universities in the surrounding area. Remedial education may also be an important function of these institutions.

Institution B enrolls just under 3,400 students and is located on the fringe of a mid-size city. With a low tuition and a combination of both day and night classes, this school attracts both part- and full-time students to its campus. Unlike Institution A, Institution B offers many more degree programs than it does certificate-granting programs. Students may complete an Associate

---

[18]A "connector" institution is defined in this report as an institution that offers some component of general education that can facilitate transfer to 4-year institutions. The institutions in the large category, *Community Connector Institutions*, are quite similar with respect to the characteristics used in this report, and therefore breaking this category down further (either by 12-month un-duplicated headcount or one of the other top variables from the cluster analysis) did not create distinguishable categories.

of Science degree in general studies, an Associate of Arts degree in liberal arts, or an Associate in Applied Science degree in specific career fields such as banking and financing, computer information systems, and nursing. In addition, there are several areas of concentration in which a student may receive a certificate. Because administrators at this institution are aware of the high percentage of their students who intend to continue their education in the 4-year college system, students are kept well-informed of which courses they can take and transfer to other public institutions without losing credit.

Community Mega-Connector Institutions (N=251)

The *Community Mega-Connector Institutions* are institutions with at least 10,000 students enrolled. They tend to be in more urban locations than institutions in the other categories of public institutions (median=2/mid-size city). These institutions have the highest median percentage of part-time students enrolled (69 percent) and the lowest median percentage of full-time, first-time degree-seeking students enrolled (7 percent) of the public institution categories. *Community Mega-Connector Institutions* grant a similar median percentage of their awards as certificates (23 percent) as do the *Community Connector Institutions*, but a lower median percentage when compared to the *Community Development and Career Institutions*. *Community Mega-Connectors* also grant a lower median percentage of their awards in occupationally specific programs (60 percent) compared to the other two groups of public institutions. Finally, this group of institutions tends to have the highest median percentage of part-time faculty (69 percent)—though it is only slightly higher when compared to the *Community Connector Institutions*.

A further look at the institutions in this category and the data described above implies that these schools are predominantly urban institutions that tend to confer awards and degrees that target job and career skills development, and to offer academic programs with some component of general education that can facilitate transfer to 4-year institutions. These institutions, which often offer a wide array of programs, resemble a 4-year university more than many of the other community colleges. According to Schuyler's research that identifies a relationship between enrollment and curriculum, these larger institutions tend to offer more liberal arts courses (Schuyler 1999). Of the awards they grant, a relatively higher percentage are degrees and are in non-occupationally specific programs. However, it is important to note that a high percentage of the full-time, first-time student body is non-degree-seeking. They are often known for providing educational services at night or on the weekend as these schools often enroll high percentages of part-time students and students over the age of 24. Finally, these large institutions, most often found in urban areas, frequently have more than one campus in the city.

An example of a *Community Mega-Connector* is Institution C, located in a city and enrolling well over 14,000 students. This institution offers university transfer and occupational training with more than 50 diverse degree and certificate areas—programs range from accounting, bricklaying, information technology, dental assistance, gambling dealer, legal assistance, to wildland firefighting. Courses are taught on evenings, weekends, and weekdays at four different locations in the area, as well as over the Internet, in an effort to be as accommodating as possible. Institution C prides itself on providing life-long learning for all those in the community.

The public institution categories do not differ greatly with respect to the percent of E&G expenditures for instruction and the percent of students enrolled that are non-traditional aged.

Private Not-For-Profit 2-Year Institutions

The percentage of total awards granted that are in allied health programs is the distinguishing characteristic of private not-for-profit, 2-year institutions (table 6; figure 5) Based on the frequency distributions, crosstabulations with other variables, and scatterplots, the percentage of total awards granted in allied health programs stood out as an obvious way for the private not-for-profit institutions to be separated. Analysis of the descriptive statistics of the characteristics of the categories, indicated that further categorization of this group of institutions was not necessary. Using this variable, two distinguishable categories of private not-for-profit institutions were formed and then named based on their characteristics as derived from both the IPEDS data and a review of the actual institutions in the category. The final categories for this sector are: *Allied Health Institutions* and *Connector Institutions* (table 7). Each category is described below and an example of an institution in each category (figure 6) is discussed.

Allied Health Institutions (N=165)

By definition, *Allied Health Institutions* grant 100 percent of their awards in allied health programs.[19] These institutions tend to be located in urban areas (median=2/mid-size city). Compared to *Connector Institutions*—the other category of private not-for-profit institutions—*Allied Health Institutions* tend to have much smaller enrollments (median=69), a much lower median percentage of part-time students enrolled (0 percent), and a lower median percentage of minority students enrolled (9 percent). On the other hand, *Allied Health Institutions* have a higher median percentage of full-time, first-time degree-seeking students enrolled (50 percent) and a higher median percentage of non-traditional aged students enrolled (52 percent). Because allied health

---

[19]For policy purposes, less extreme cut-off points for forming categories may be appropriate. For ease of analysis, and lack of analytic support for other potential cut-off points, 100 percent of awards in allied health fields was chosen for placing these institutions into categories.

Table 6.—Descriptive statistics of the percentage of degrees awarded in allied health programs at private not-for-profit, 2-year institutions

N	
Valid	293.0
Missing	16.0
Mean	61.0
Median	100.0
Standard Deviation	46.5
Variance	2,159.8
Range	100.0
Minimum	0.0
Maximum	100.0
Percentiles	
25%	0.0
50%	100.0
75%	100.0

SOURCE: U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), Full 1997 Collection Year.

Figure 5.—Frequency distribution of private not-for-profit, 2-year institutions by the percentage of awards granted in allied health programs

Percentage of degrees awarded in allied health programs
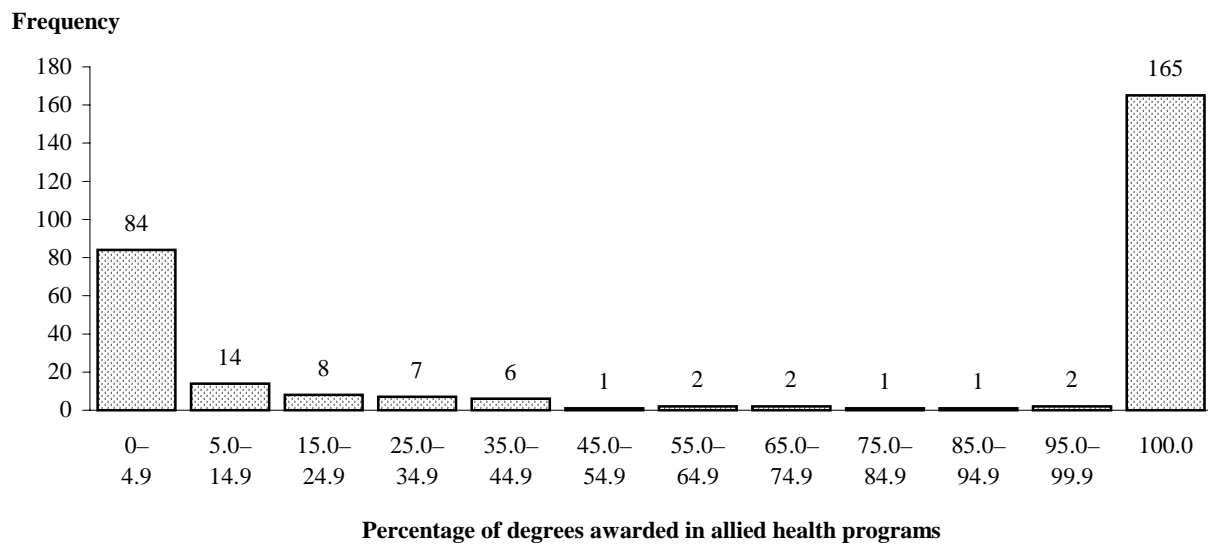
SOURCE: U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), 1997 Full Collection Year.

Table 7.—Characteristics of the categories for private not-for-profit, 2-year institutions, according to selected variables

	Allied Health N= 165		Connector N=128	
	Mean	Median	Mean	Median
Urbanicity	—	2.0	—	2.0
12-month unduplicated headcount	88.4	69.0	649.7	435.5
Percentage of full-time, first-time degree-seeking students	50.4	48.1	39.0	37.5
Percentage of part-time students	15.5	0.0	21.0	13.2
Percentage of minority students	15.8	9.2	28.2	15.0
Percentage of non-traditional aged students	56.0	52.2	35.8	37.0
Percentage of part-time faculty	20.2	7.0	46.7	45.8
Percentage of awards granted as certificates	88.0	100.0	29.9	7.5
Percentage of awards granted in allied health programs	100.0	100.0	10.8	0.0
Percentage of awards granted in occupationally specific programs	100.0	100.0	65.5	88.8
Percentage of E&G expenditures for instruction	64.1	66.7	29.0	27.6

— Not applicable.

NOTE: The sum of the number of institutions in each category does not add to the total number of private, not-for-profit institutions due to missing data in the variables chosen for categorization.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), Full 1997 Collection Year.

programs are, by definition, occupationally specific programs, 100 percent of awards granted by *Allied Health Institutions* are in occupationally specific programs. In addition, these institutions have a high median percentage of awards that are granted as certificates (100 percent). Finally, *Allied Health Institutions* have a much lower median percentage of part-time faculty (7 percent) while having a much higher median percentage of E&G expenditures for instruction (67 percent).

Institution D, an *Allied Health Institution*, uses the facilities of a local community teaching hospital. The school offers a certificate in nursing program, which prepares students to take the licensing exam to become a registered nurse. The stated mission of this school is to provide a program that assists its graduates in acquiring the competencies needed for the practice of professional nursing. Located in a large city, this commuter institution only enrolls approximately 50 students, all of whom attend full-time.

Connector Institutions (N=128)

Connector Institutions are those private not-for-profit institutions that do not grant 100 percent of their awards in allied health programs. These institutions tend to be located in urban areas (median=2/mid-size city). The institutions in this category also tend to have larger enrollments

Figure 6.—Characteristics of the private not-for-profit, 2-year institution examples, according to selected variables
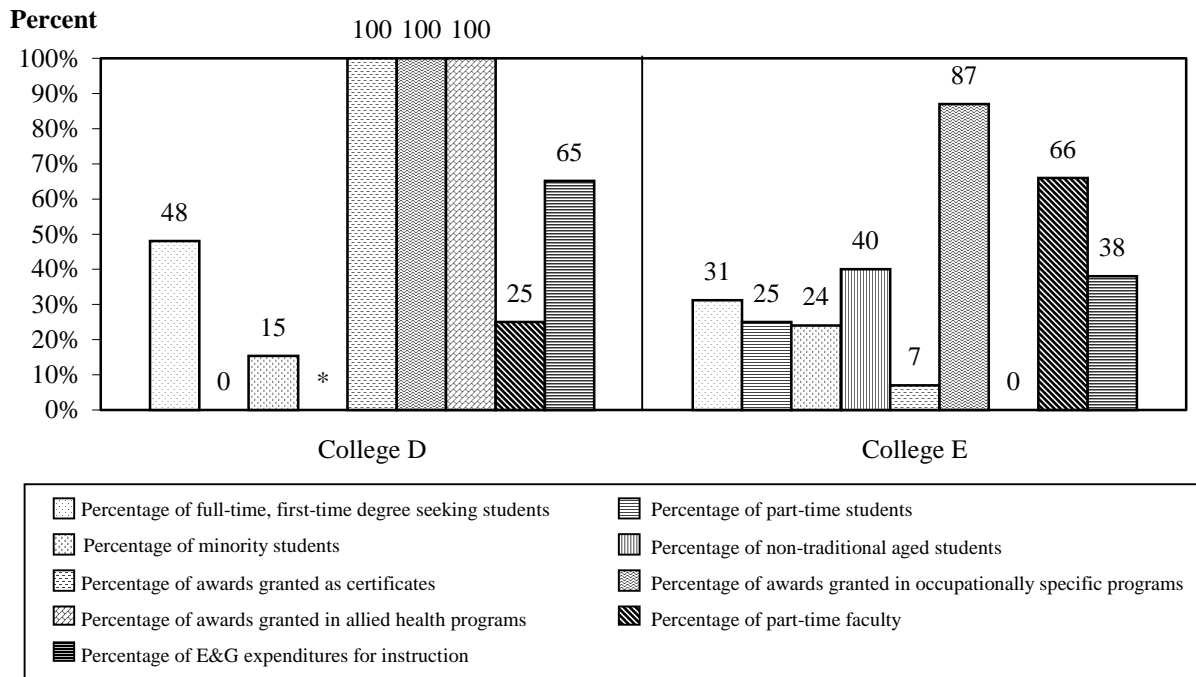
*Indicates the data is not available.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), 1997 Full Collection Year.

(median=436), a higher median percentage of part-time students (13 percent) and a lower median percentage of full-time, first-time degree-seeking students (38 percent) than *Allied Health Institutions*. *Connector Institutions* have a lower median percentage of non-traditional aged students (37 percent), but a higher median percentage of minority students enrolled (15 percent) than *Allied Health Institutions*. These institutions have a low median percentage of certificates awarded (8 percent), and though the percentage of awards granted in occupationally specific programs is less than that of the *Allied Health Institutions*, *Connector Institutions* still tend to confer a high percentage of their awards in occupationally specific fields (median=89 percent). Finally, these institutions have a higher median percentage of part-time faculty (46 percent), but a much lower median percentage of E&G expenditures for instruction (28 percent).

From the data presented and from examining the institutions in this category, *Connector Institutions* seem to be larger than the *Allied Health Institutions,* however these institutions are still small in enrollment size. *Connector Institutions* are schools that tend to confer awards and degree that target job and career skills development, and to offer academic programs with some component of general education that can facilitate transfer to 4-year institutions for those students who

wish to continue their education. This category consists of a mix of institutions, many of which confer all of their awards in occupationally specific programs. Depending on their interests, researchers may consider further dividing this small group of institutions by the percentage of awards granted in occupationally specific programs.

Institution E, with an enrollment of just under 500 full- and part-time students, is a typical *Connector Institution*. Students of Institution E can earn an Associate of Arts, Associate of Science, or an Associate of Applied Science degree. Graduates are then prepared either to enter or return to the workforce or transfer to a 4-year college. The curriculum for each program includes a liberal arts core of courses and courses in one of the 15 majors that are offered. Institution E is proud of its mix of traditional and non-traditional students and believes that it offers all students small classes and personal attention, allowing for both academic and vocational needs to be addressed.

Private For-Profit 2-Year Institutions

Virtually all institutions in this sector are found in urban locations and grant a high percentage of their awards in occupationally specific programs. These variables were not included in this particular sub-analysis because they are nearly universal and do not discriminate. The percentage of total awards granted that are certificates is a distinguishing characteristic of private for-profit, 2-year institutions (table 8; figure 7). Based on cluster analysis and subsequent analyses on the driving variables within this sector of institutions, the percentage of total awards granted that are certificates creates two distinguishable categories that were then named based on IPEDS data and a review of actual institutions in the respective categories: *Career Connector Institutions* and *Certificate Institutions* (table 9). In addition to descriptions of the categories, examples of institutions in each category are presented (figure 8).

Career Connector Institutions (N=367)

This category consists of institutions that grant less than 100 percent of their awards as certificates—at least one of their awards conferred is in the form of an associate's degree.[20] The median percentage of awards granted as certificates is 28 percent. The median enrollment at institutions in this category is 391 students—higher than the median enrollment for *Certificate Institutions*. *Career Connector Institutions* have a lower median of full-time, first-time degree seeking students (51 percent), but a higher median percentage of part-time faculty on staff (48

---

[20]For policy purposes, other, less extreme, cut-off points for forming categories could be appropriate. For ease of analysis in this report, and lack of analytic support for other potential cut-off points, 100 percent of awards granted as certificates was chosen as the cut-off point for placing these institutions into categories.

Table 8.—Descriptive statistics of the percentage of certificates awarded at private for-profit, 2-year institutions

N	
Valid	700.0
Missing	30.0
Mean	65.6
Median	95.9
Standard Deviation	40.2
Variance	1,618.0
Range	100.0
Minimum	0.0
Maximum	100.0
Percentiles	
25%	25.6
50%	95.9
75%	100.0

SOURCE: U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), Full 1997 Collection Year.

Figure 7.—Frequency distribution of private for-profit, 2-year institutions by the percentage of awards granted as certificates
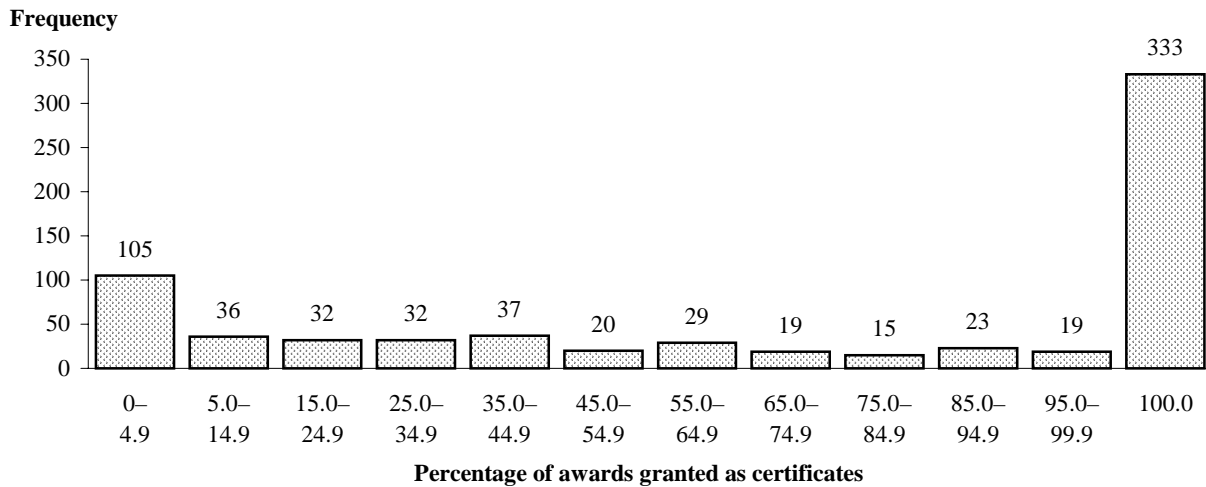
SOURCE: U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), 1997 Full Collection Year.

Table 9.—Characteristics of the categories for private for-profit, 2-year institutions, according to selected variables

	Career Connector N=367		Certificate N=333	
	Mean	Median	Mean	Median
Urbanicity	—	2.0	—	2.0
12-month unduplicated headcount	609.3	391.0	204.5	102.0
Percentage of full-time, first-time degree-seeking students enrolled	53.0	50.9	65.2	76.9
Percentage of minority students	30.2	25.7	29.1	19.2
Percentage of non-traditional aged students enrolled	42.4	42.0	43.3	46.0
Percentage of awards granted as certificates	34.4	27.7	100.0	100.0
Percentage of part-time faculty	43.4	48.3	29.4	25.0
Percentage of E&G expenditures for instruction	29.7	25.0	34.0	30.1

— Not applicable.

NOTE: The sum of the number of institutions in each category does not add to the total number of private, for-profit institutions due to missing data in the variables chosen for categorization.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), Full 1997 Collection Year.

Figure 8.—Characteristics of the private for-profit institution examples, according to selected variables
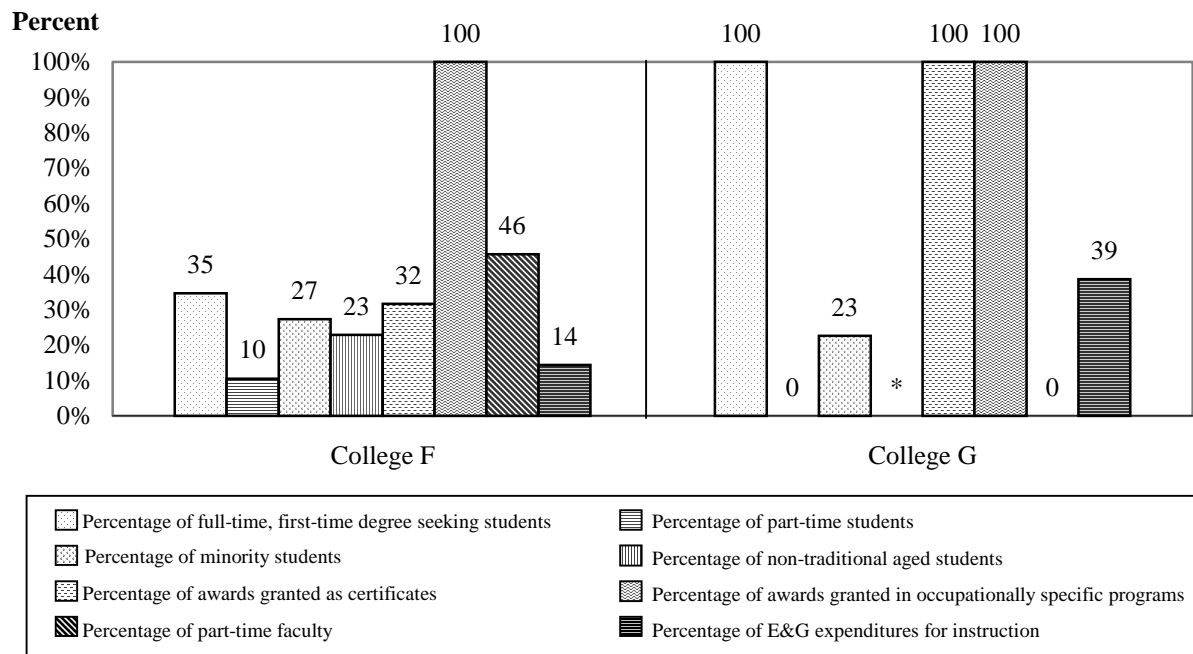
*Indicates the data is not available.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), 1997 Full Collection Year.

percent) than do *Certificate Institutions*. The median location for institutions in this category is a "mid-size city" (median=2).

The private for-profit *Career Connector Institutions* are degree-granting institutions—although many also offer certificates—that target job and career skills development for students. Many of these institutions offer academic programs with some component of general education that can facilitate transfer to 4-year institutions for those students who wish to continue their education. Located almost exclusively in urban areas, these relatively small schools concentrate their program offerings in occupations that are in high demand, particularly technology and business. A larger percentage of their students are minority.

Institution F grants associate's degrees and certificates in order to provide students career education for entry-level positions in business, allied health, and travel and tourism. Located in a city, Institution F enrolls approximately 920 part- and full-time students. Institution F is owned by an educational corporation in the area and places great emphasis on guiding students through the process of securing and maintaining employment.

Certificate Institutions (N=333)

Certificate Institutions are institutions that grant 100 percent of their awards as certificates, as opposed to associate's degrees.[21] The institutions in this category tend to be located in more urban locations—they are most often found in a "mid-size city" (median=2). As a group, they tend to have smaller enrollments than do the *Career Connector Institutions*—the median enrollment is 102 students. *Certificate Institutions* have a high median percentage of full-time, first-time degree seeking students (77 percent), when compared to the *Career Connector Institutions*. Finally, these institutions have a rather low median percentage of part-time faculty (25 percent).

Certificate Institutions tend to give specialized training usually in a single job category or area, preparing students to proceed directly into the workforce. The vast majority are located in an urban area and enroll many full-time students with the goal of obtaining a certificate or award. This category includes a high concentration of schools focusing on cosmetology services. Many of the programs offered at these institutions require state certification.

Institution G is one branch of a cosmetology consortium in the state in which it is located. Like the other branches, it is an urban institution. This particular site enrolls 72 full-time stu-

---

[21]It is important to that the institutions in this study must offer at least one program (regardless of whether it is a degree-granting or certificate-granting program) of at least 2-years but less than 4 academic years in duration to be included in the analysis. All institutions in this study are considered 2-year institutions according to 1997 IPEDS data; less-than-2-year schools were not included.

dents. Certificates are offered in a range of cosmetology programs, each requiring a specific number of class hours in order to earn the diploma. After completing the program, students are prepared to take the state licensing examination in their program area.

Figure 9.—Classification system of 2-year institutions
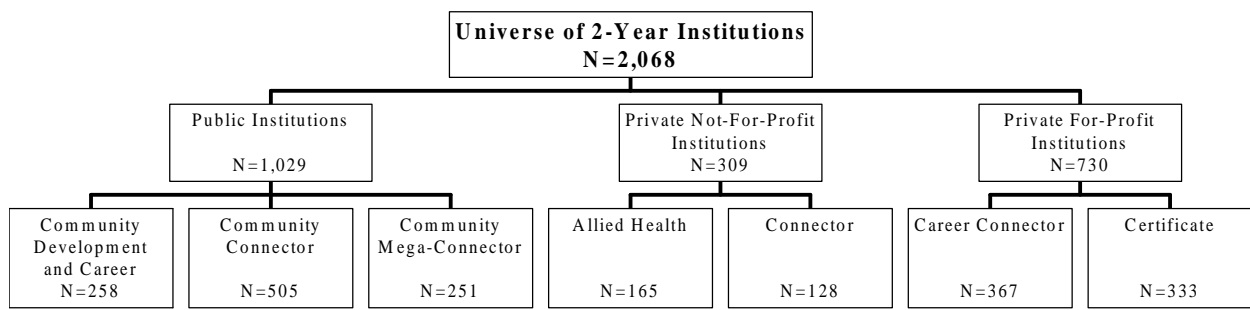
```
                    ┌─────────────────────────────┐
                    │ Universe of 2-Year Institutions │
                    │          N=2,068             │
                    └─────────────────────────────┘
        ┌──────────────────────┼──────────────────────┐
┌────────────────┐  ┌────────────────────┐  ┌────────────────────┐
│ Public Institutions │  │ Private Not-For-Profit │  │  Private For-Profit  │
│                │  │     Institutions    │  │     Institutions    │
│   N=1,029      │  │       N=309         │  │       N=730         │
└────────────────┘  └────────────────────┘  └────────────────────┘
```

| Community Development and Career N=258 | Community Connector N=505 | Community Mega-Connector N=251 | Allied Health N=165 | Connector N=128 | Career Connector N=367 | Certificate N=333 |

NOTE: The sum of the number of institutions in each category does not add to the total number of institutions due to missing data in the variables chosen for categorization. In the universe of 2,068 institutions analyzed in this report, 61 institutions could not be placed in a final category: 15 public 2-year institutions, 16 private not-for-profit institutions, and 30 private for-profit institutions.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), 1997 Full Collection Year.

THIS PAGE INTENTIONALLY LEFT BLANK

Summary

Originally published in 1973 by the Carnegie Commission on Higher Education, The Carnegie classification has made significant contributions to research and policy in higher education. While the Carnegie system is widely accepted, it is often criticized for its hierarchical nature, and has mistakenly come to be seen as a kind of ranking system for colleges. This misinterpretation is one of the reasons the Carnegie Foundation for the Advancement of Teaching is in the process of revising its classification system—a preliminary revision was released in the fall of 2000, with the final revision scheduled to be completed by 2005.

A limitation of the present Carnegie classification is that all 2-year institutions are placed into a single category of over 1,600 institutions. Not only is this the largest category by far, but the diversity within this large group of institutions is not captured. In addition, the Carnegie classification includes only accredited and degree-granting institutions, thereby excluding approximately 700 existing 2-year institutions. The classification system offered by Katsinas classifies a large universe of 2-year institutions—2,501 institutions—and uses nationally available data to do so. However, the techniques used to arrive at the classification system are unclear, making it difficult to verify and replicate. A strength of the system proposed by Schuyler is that it captures the institutional mission at community colleges by collecting course catalogue data. However, the use of course catalogues bases the system on data that is not nationally available nor can be easily accessed. In addition, the Schuyler system only categorizes public Associate of Arts colleges. Similarly, though the NCPI market based classification system is clear and well-defined, its focuses mostly on public 2-year institutions.

Thus, the impetus behind the development of a new classification system for 2-year institutions rests upon the fact that the present systems—all with their own strengths—do not fully and adequately describe this vital sector of the postsecondary community. The goal in creating this classification system is to provide a new framework for policy discussions and a basis for research and analysis. As such, its development had to satisfy several criteria; the following is a list of the criteria, accompanied by a description of how each was addressed.

The data must be collected at the national level. The data source was the National Center for Education Statistics' Integrated Postsecondary Education System (IPEDS). This comprehensive system collects data on enrollment, program completion, faculty, staff, finances, and academic libraries from all postsecondary institutions.

Each variable, and the classification system as a whole, must be replicable so that the classification system will continue to be useful in the future. Since the variables used in the classification system are all derived from IPEDS, they are available to anyone who wishes to analyze them. Furthermore, the variables used are collected on a consistent basis.

The classification system must have a manageable number of categories and each institution should reside in one, and only one, category. The classification scheme resulted in seven mutually exclusive categories; three categories in the public sector; two categories in the private not-for-profit sector; and two categories in the private for-profit sector.

The classification system must make policy-relevant distinctions among institutions. The variables chosen from the IPEDS data collection system are useful with regard to policy development and institutional research. Some examples of these variables include student enrollment, percentage of full-time, first-time degree-seeking undergraduates, percentage of part-time students, and percentage of minority students. However, it cannot be overlooked that many policy relevant data elements central to the mission of 2-year institutions are not collected at the national level. This limitation notwithstanding, several variables did relate to the functions of 2-year institutions and were, therefore, relevant to research and policy analysis.

The categories must be descriptive and not evaluative or hierarchical in nature. The variables employed to categorize institutions are those that are commonly used and value neutral. Care was taken to choose category names that did not suggest a hierarchy so that institutions within a particular category will not feel pressured to "move up" to another category because of a perceived higher ranking.

The resulting seven-category classification system is presented below.

Public Institutions

Community Development and Career Institutions are institutions with an unduplicated headcount of less than 2,000 students. These institutions tend to confer awards and degrees primarily in job and career skills development and to focus on overall workforce development for the communities they serve.

Community Connector Institutions are institutions with an unduplicated headcount of 2,000 to 9,999 students. These institutions tend to confer awards and degrees that target job and career skills development, and to offer academic programs with some component of general education that can facilitate transfer to 4-year institutions.

Community Mega-Connector Institutions are institutions with an unduplicated headcount of at least 10,000 students. These institutions tend to be in urban locations, to confer awards and degrees that target job and career skills development, and to offer academic programs with some component of general education that can facilitate transfer to 4-year institutions.

Private Not-For-Profit Institutions

Allied Health Institutions are institutions that grant 100 percent of their awards in Allied Health programs. These institutions tend to be small in enrollment and to have an exclusive focus on Allied Health training.

Connector Institutions are institutions that grant less than 100 percent of their awards in Allied Health programs. These institutions tend to confer awards and degrees targeting job and career skills development, and to offer academic programs with some component of general education that can facilitate transfer to 4-year institutions.

Private For-Profit Institutions

Career Connector Institutions are institutions that grant less than 100 percent of their awards as certificates. They are degree-granting institutions—although many also offer certificates—that target job and career skills development. Many of these institutions offer academic programs with some component of general education that can facilitate transfer to 4-year institutions.

Certificate Institutions are institutions that grant 100 percent of their awards as certificates. These institutions provide specialized training, usually in a single job category or area.

THIS PAGE INTENTIONALLY LEFT BLANK

References

Aldenderfer, M.S., and Blashfield, R.K. (1984). *Cluster Analysis*. Newbury Park, CA: Sage Publications, Inc.

Anderberg, M.R. (1973). *Cluster Analysis for Applications*. New York: Academic Press.

The Carnegie Foundation for the Advancement of Teaching. (2000). *The Carnegie Classification of Institutions of Higher Education*. Available: *http://www.carnegiefoundation.org/ Classification/CIHE2000/Tables.htm* [May 31, 2001].

Katsinas, S.G. (1993). *Towards a Classification System for Community Colleges.* Paper presented at the Annual Meeting of the Council of Colleges and Universities, Portland, OR.

Katsinas, S.G. (1996). Preparing Leaders for Diverse Institutional Settings. *New Directions for Community Colleges, 95*: 93–98.

McCormick, A.C. (1999). *Plans for the 2000 Edition of the Carnegie Classification*. Presentation to the Washington Higher Education Secretariat.

National Center for Postsecondary Improvement. (1998). The User-Friendly Terrain: Defining the Market Taxonomy for Two-Year Institutions. *Change*: 35–38.

Schuyler, G.L. (1999). *A Curriculum-Based Classification System for Community Colleges*. Unpublished doctoral dissertation.

SPSS, Inc. (1999). *SPSS Base 10.0 Applications Guide*. United States: Author.

THIS PAGE INTENTIONALLY LEFT BLANK

Appendix A—Technical Notes

The Integrated Postsecondary Education Data System

The Integrated Postsecondary Education Data System (IPEDS) is a comprehensive database that encompasses all identified institutions whose primary purpose is to provide postsecondary education. IPEDS consists of institutional-level data that can be used to describe trends in higher education at the institutional, state, and/or national levels.

Postsecondary education is defined within IPEDS as the provision of formal instructional programs whose curriculum is designed primarily for students who have completed the requirements for a high school diploma or its equivalent. This includes academic, vocational, and continuing professional education programs, and excludes avocational and adult basic education programs.

IPEDS includes information about baccalaureate or higher degree-granting institutions, 2-year award institutions, and less-than-2-year institutions (i.e. institutions whose awards usually result in terminal occupational awards or are creditable toward a formal 2-year or higher award). Each of these three categories is further disaggregated by control (public; private not-for-profit; and private for-profit), resulting in nine institutional categories or sectors.

Specialized, but compatible, reporting formats have been developed for these nine sectors of postsecondary education providers. In general, the surveys/reports developed for postsecondary institutions granting baccalaureate and higher degrees are the most extensive; forms for the 2-year and less-than-2-year awards granting sectors request less data. This design feature accommodates the varied operating characteristics, program offerings, and reporting capabilities that differentiate postsecondary institutional sectors while yielding comparable statistics for all sectors.

Data are collected from approximately 11,000 postsecondary institutions. IPEDS has been designed to produce national-, state-, and institutional-level data for most postsecondary institutions. However, prior to 1993, only national-level estimates from a sample of institutions were available for the private, less-than-2-year institutions.

Data in IPEDS have been organized into several survey areas. For this report, the most important surveys include the following[22]:

Institutional Characteristics, including institutional names, addresses; congressional districts; counties; telephone numbers; tuition; control or affiliation; calendar systems; levels of degrees and awards offered; types of programs; and accreditation for all postsecondary education institutions in the United States and outlying territories.

Fall Enrollment, including information about full- and part-time enrollment by racial/ethnic category and sex for undergraduates, first-professional, and graduate students. Age distributions by level of enrollment and sex are collected in odd-numbered years, and first-time degree-seeking student enrollments by residence status are collected in even-numbered years.

Completions, including annual counts of associate's, bachelor's, master's, doctor's, and first-professional degrees and other formal awards, by 6-digit Classification of Instructional Program (CIP) code, race/ethnicity, and sex of recipient. Prior to 1994-95, awards by race/ethnicity were only collected by the 2-digit program area.

Financial Statistics, including each institution's current fund revenues by source (e.g., tuition and fees, government, gifts); current fund expenditures by function (e.g., instruction, research); assets and indebtedness; and endowment investments.

Fall Staff, including the number of staff by occupational activity, full- and part-time status, sex, and race/ethnicity. Data are collected in odd-numbered years. Beginning with 1993, this survey replaces the EEO-6 survey conducted by the Equal Employment Opportunity Commission.

Other components of IPEDS have provided data on faculty salaries and tenure, and library resources.

Selecting the Universe of Institutions

The universe of institutions in the analysis is based on the 1997–98 Institutional Characteristics Survey. First, 2-year institutions with Title IV participation agreements were selected from the total IPEDS universe of 2,800 2-year institutions. In all, 373 institutions were excluded be-

---

[22]The descriptions of the IPEDS surveys that are provided are valid for IPEDS data collected from 1993–1999, which includes the data that was used in this report. It is important to note, however, that this information is not necessarily valid for the more recently released web-based collection procedure. Data items that were collected in the past (including the years of data used for this report) may no longer be included in the IPEDS surveys. Detailed information about IPEDS is available at the National Center for Education Statistics Web site (*http://nces.ed.gov*).

cause they were not eligible for Title IV student aid funds—32 public, 207 private not-for-profit, and 134 private for-profit.

The universe was then further trimmed based on data reporting. Institutions had to either have reported all their data as an individual institution or reported consistently across the surveys as either a parent or child institution to be included in the analysis. A parent institution reports combined data, including another institution's data (the child institution) with their own. For example, an institution that is part of a state college system may report data individually or in combination with the other institutions in the system. Often institutions report their data inconsistently across surveys—for one survey they will report in combination with another institution (as a parent or child) while in a different survey they may report their data individually. This makes analysis with these institutions difficult, therefore institutions that reported inconsistently were not included in the universe for this analysis. In total 359 institutions were excluded due to reporting their data inconsistently as a parent or child across the IPEDS surveys, or due to excessive missing data in the key variables for this analysis—219 public, 39 private not-for profit, and 101 private for-profit institutions. The remaining 2,068 2-year institutions—all with Title IV participation agreements—were included in the analysis.

Accuracy of Estimates

The statistics in this report are estimates derived from a population. Because this report uses a census of an entire population there is not a sampling error however, there is still the possibility of nonsampling error.

Nonsampling errors can be attributed to a number of sources: inability to obtain complete information from all institutions (i.e. some institutions refused to participate, or participated but answered only certain items); ambiguous definitions; differences in interpreting questions; inability or willingness to give correct information; mistakes in recording or coding data; and other errors of collecting, processing, sampling and imputing missing data.

The estimates produced in this report were produced using the SPSS 10.0 software package. SPSS makes it possible for users to specify and generate a variety of univariate and multivariate analyses, including several types of cluster analysis.

Statistical Procedures

The primary statistical procedure used in this report is K-means cluster analysis. Cluster analysis is a multivariate procedure for detecting groupings in data and is considered a good

technique to use in exploratory data analysis, when the data sample is suspected to not be ho-
mogenous. Because K-means cluster analysis requires specifying the number of clusters to be
formed, a technique to determine an optimal number of clusters was performed—hierarchical
cluster analysis.

The hierarchical cluster analysis procedure uses an algorithm that starts with each case in a
separate cluster and combines pairs of clusters, or a single case, with a cluster until all cases are
joined into one cluster. The method is hierarchical because once two cases or cluster are joined,
they remain together until the final step—a cluster formed in a later stage of the analysis contains
clusters from an earlier stage and so on.

Hierarchical cluster analysis was run on six different random samples of the universe of in-
stitutions—the size of the samples ranged from approximately 100 to 300 institutions. In SPSS—
the software package used—hierarchical clustering excludes all cases with values missing for any
of the variables used in the analysis. All variables included in the cluster analyses first were stan-
dardized into z scores because the variables are measured on different scales. The linking method
used was average linkage (also known as between-groups linkage) and the distance measure was
the squared Euclidean distance, defined as the sum of the squared distances over all variables.
The squared Euclidean distance is the most frequently used distance measure for clustering cases,
defined as

$$d_{ij} = \sum_{k=1}^{p} (x_{ik} - x_{jk})^2$$

where d_{ij} is the distance between cases i and j, and x_{ik} is the value of the kth variable for the ith
case. Two cases are identical when the distance between them is zero.

Hierarchical clustering computes a distance matrix with entries for every pair of cases
(which is why a large number of cases becomes unwieldy) as well as icicle plots and dendo-
grams, that provide a picture of when each case is joined with another. From these graphical dis-
plays and distance measures, a judgment can be made about the number of clusters that is
appropriate and therefore should be specified in a K-means analysis. Smaller distances indicate
that fairly homogenous clusters have been joined, while larger distances are an indication that the
members of the clusters joined are more dissimilar. When there is a sudden jump in the size of
this distance (or fusion coefficient), it should be considered that a solution is reached at that point
and that the procedure has reached an optimal number of clusters. From the hierarchical cluster
analysis conducted on the samples from the universe of institutions used in this report, it was de-
termined that between 6 and 8 clusters were an appropriate number of categories.

In K-means cluster analysis, the procedure begins by using the values of the first k cases in the data file as temporary estimates of the k cluster means, where k is the number of clusters specified by the user (6, 7, and 8 clusters, for the purposes of this report). Initial cluster centers are formed by assigning each case to the cluster with the closest center (using the squared Euclidean distance measure described above) and then updating the center. An iterative process is used to find the final cluster centers. At each step, cases are grouped into the cluster with the closest center, and the cluster centers are recomputed. This process continues until no further changes occur in the centers or until the maximum number of iterations is reached.

As with the hierarchical analysis, all variables included in the cluster analysis were standardized into z scores prior to the procedure. In the K-means procedure, cases were "excluded pairwise," meaning cases were assigned to clusters based on distances computed from all variables with non-missing values. The maximum number of iterations was set at the maximum for SPSS, 999.

The output from the analysis provides the distance from each case to its cluster center, characterizing whether or not a case is close to the others within its cluster or is an outlier. In addition, for each variable, individually, SPSS computes a one-way analysis of variance using the final clusters as groups. The between-cluster mean square and the within-cluster mean square are displayed, and the ratio of these two mean squares is the ANOVA F statistic. The size of the F statistic is useful for identifying variables that drive the clustering and also those that differ little across the clusters. However, because the clusters are formed to characterize differences, the significance levels of the F ratios should be ignored.

There are a few cautions about the cluster analysis procedures that should be noted. First, the strategy of cluster analysis is "structure-seeking" although its operation is "structure-imposing." Clustering methods are used to discover structure in data, and will always place objects into groups, whether or not the groups are "real," "natural," and/or optimal, or if the groups are simply imposed on the data by the method. In addition, groups can be radically different in composition when different clustering methods are used (Aldenderfer and Blashfield 1984). Finally, the K-means method of cluster analysis is very sensitive to poor initial partitions, and this problem is exacerbated by the selection of a random initial partition—a popular option with much iterative software, including SPSS.

THIS PAGE INTENTIONALLY LEFT BLANK

Appendix B—Glossary

This glossary describes the variables used in this report, which come from FINALDATA.SAV, an analysis file created in SPSS from IPEDS data. These variables were either items taken directly from the IPEDS surveys or they were derived by combining one or more items in these surveys. For all of the variables in this glossary, the variable name contained in the analysis file is identified in the right-hand column.

Institutional characteristics are listed first, followed by fall enrollment, completions, fall staff, and financial statistics. Variables are listed alphabetically by the variable name, under the appropriate IPEDS survey.

GLOSSARY INDEX

Institutional Characteristics

Carnegie classification code   **CARNEGIE**

Indicates the Carnegie classification of the reporting institution. Only the Carnegie categories relevant to this report are listed below.

Associate of arts colleges = 40	These institutions offer associate of arts certificate or degree programs and, with few exceptions, offer no baccalaureate degrees.
Health profession schools = 53	The specialized institutions in this category award most of their degrees in such fields as chiropractic, nursing, pharmacy, or podiatry.
Tribal colleges = 60	These specialized institutions are, with few exceptions, tribally controlled and located on reservations. They are all members of the American Indian Higher Education Consortium.

Not classified in the Carnegie classification
  system = 0

Percent of total undergraduates who were full-time, first-time degree seeking   **FTFTPERC**

The percentage of total undergraduates (unduplicated) composed of full-time, first-time degree seeking undergraduates.

Highest degree offered   **HDEGOFFR**

The highest degree offered at the reporting institution.

        Associate's degree = 40
        Non-degree granting = 0

Degree of urbanization of location of the institution   **LOCALE2**

The degree of urbanization of the reporting institution's location, based on U.S. Census Bureau designations.

Large city = 1	A central city of a CMSA or MSA having a population greater than or equal to 250,000
Mid-size city = 2	A central city of a CMSA or MSA, having a population less than 250,000
Urban fringe of a large city = 3	Any incorporated place, CDP, or non-place territory within a CMSA or MSA of a large city and defined as urban by the Census Bureau.

Urban fringe of a mid-size city = 4	Any incorporated place, CDP, or non-place territory within a CMSA or MSA of a mid-size city and defined as urban by the Census Bureau.
Large town = 5	Any incorporated place or CDP with a population greater than or equal to 25,000 and located outside a CMSA or MSA.
Small town = 6	Any incorporated place or CDP with a population less than 25,000 and greater than or equal to 2,500 and located outside a CMSA or MSA.
Rural = 7	Any incorporated place, CDP, or non-place territory designated as rural by the Census Bureau.

Sector of institution   **SECTOR**

Indicates the sector—level and control—of the reporting institution. When limited to 2-year institutions, control is the only variance between different sectors. Only the following sectors were included in this report:

Public, 2-year = 4
Private not-for-profit, 2-year = 5
Private for-profit, 2-year = 6

Unduplicated headcount of undergraduate students enrolled in a 12-month period   **TOSTUCU**

The unduplicated of all undergraduate students enrolled through the 12-month period.

Fall Enrollment

Percent of total undergraduates who are 24 years old or older   **PERCAGE**

The percentage of total undergraduates composed of 24 years old or older undergraduates at the reporting institution.

Percent of total undergraduates who are minorities   **PERCMIN**

The percentage of total undergraduates composed of minority undergraduates at the reporting institution.

Percent of total undergraduates who were part-time   **PERCPART**

The percentage of total undergraduates composed of part-time undergraduates at the reporting institution.

Completions

Percent of total awards granted that are in allied health fields   **PCTAH2**

The percentage of total awards granted that are in allied health fields. Allied Health Fields are designated by the U.S. Department of Education's *Classification of Instructional Programs* code 51 (Health Professions and Related Sciences).

Percent of total awards granted that are certificates   **PCTCERT**

The percentage of total awards granted that are certificates.

Percent of awards granted that are in occupationally specific fields   **PCTOCC2**

The percentage of total awards granted that are in occupationally specific fields. Occupationally specific fields are designated by the U.S. Department of Education's *Classification of Instructional Programs* as "instructional programs whose expressed intent is to impart work-related knowledge and skills at the secondary and postsecondary levels. The term has been historically applied to vocational programs offered in grades 11 and 12 at the secondary level, and to postsecondary vocational and technical education programs at the sub-baccalaureate level.

Fall Staff

Percent of total faculty that are part-time   **PERCFPT**

The percentage of total faculty composed of part-time faculty, at the reporting institution.

Finance

Percent of E&G expenditures that for instruction   **PCTINSTR**

The percentage of education and general expenditures at the reporting institution, that are used for instructional purposes.

Percent of total current funds revenues that are from state and local appropriations, grants, and contracts
        STLOCSUP

The percentage of total current funds revenues at the reporting institution, that are from state and local appropriations, grants and contracts.