Title: A Cluster Analysis Broadly Characterizing Engineering/Technology Management Programs

Course:
Year: 1994
Author(s): A. P. J. Sandoval

Report No: P94007

Abstract: This project is to broadly group ETMs into several clusters based on their date of establishment, name of program, academic base, and type of degrees offered. The nature of cluster analysis does not lend itself to hypothesis testing, but the results can help identify other statistical methods. Chi-Square contingency tables to test independence among pairs of variables may be possible. However, some variables are not mutually exclusive and may have valid inputs in more than one cells.

A Cluster Analysis Broadly Characterizing Engineering/Technology Management Programs

Alfredo P. Sandoval, Jr.

EMP-P9407

Engineering Management 610
Research Methodology in Engineering Management
Winter Term 1994

A Cluster Analysis Broadly Characterizing Engineering/Technology Management Programs

submitted by:
Alfredo P. Sandoval, Jr.

Engineering Management Program
Portland State University
March 11, 1994

A Cluster Analysis Broadly Characterizing Engineering/Technology Management Programs

Background

The Engineering Management Program of Portland State University is currently conducting the fifth of a series of longitudinal studies whose objective is to update the information on educational and research characteristics of Engineering/Technology Management (ETM) Programs. It is desired to identify the commonalities and differences among these programs. Both qualitative and quantitative data, as well as comments and lists are available from the responses. The field is experiencing a rapid growth and a database containing information about the identified ETM programs will enhance the field and contribute to its strategic direction.

Limitations

The survey questionnaire consists of 44 questions which address the educational and research characteristics of ETMs. The cluster analysis that will be performed will be the first attempt to use multivariate analysis apart from the descriptive statistics used before. Based on the responses received so far, it is evident that the structure of international ETMs are quite different from that of American ETMs. To avoid misrepresentation, only responses from American institutions will be considered.

The dataset used consisted of 97 responses from contact persons identified in the initial inquiry forms which asked if their institution has an ETM program. This number represents approximately 70% of the total responded expected. Observations indicate that programs have different names, academic bases, degrees offered, and a substantial range of dates of

In order to make it visually easier to determine the number of clusters to be extracted (at least in the mathematical sense), dendograms were requested. This hierarchical clustering solution shows the step by step clustering solution, from single entities to everything belonging to one cluster. Subjective judgments on the number of clusters considered based on the dendogram display are tested for Group Mean Significance through the use of the default QUICK CLUSTER command, requesting ANOVA. A specified number of clusters to be created is chosen and requested. Final classification cluster centers are identified and levels of significance comparing the difference between the group means are provided. Several smaller models which do not use all the variables were also tried, using the COMPLETE METHOD.

The list of the commands used in the analysis are as follows:

DESCRIPTIVES VARIABLES EMGT MOT AB1 AB2 AB3 BS MS PHD
/STATISTICS 12 - sums for each category are requested
DESCRIPTIVES VARIABLES DATE (ZDATE)
/STATISTICS 1 10 11 - requests a z-transformation, mean, minimum, and maximum (see [2])
CLUSTER EMGT MOT AB1 AB2 AB3 BS MS PHD ZDATE
/PLOT DENDOGRAM - requests the default cluster analysis (hierarchical) and the dendogram

/METHOD = COMPLETE - requests the Complete cluster method of combining clsuters (see [3])

QUICK CLUSTER EMGT MOT AB1 AB2 AB3 BS MS PHD ZDATE
/CRITERIA CLUSTERS (6) /PRINT CLUSTER DISTANCE ANOVA
- requests 6 clusters using the non-hierarchical clustering procedure; initial seeds are randomly selected from the observations [2,p.278]; the ANOVA table is then provided. (see [3])

Results

Several cluster analyses using a subset of the four variables were conducted. Using DATE only (E1) does not show meaningful clusters. As expected, using ZDATE produced the exact dendogram. ZDATE, EMGT, and MOT were used next (E2); the results revealing minimum clustering. The ANOVA from the QUICK CLUSTER procedure showed that just about any number of clusters can be defined if the significant differences between the cluster means are considered. A run (E3) using EMGT, MOT, AB1, AB2, AB3, BS, MS, & PHD showed several hierarchies of clusters. All variables, with DATE (unscaled) revealed a dendogram resembling the run with only DATE considered. This was expected as the years have a very large relative magnitude when compared to binary values of zero and one. All these runs except QUICK CLUSTER used METHOD = COMPLETE.

Finally, two runs using all the available information (ZDATE was used) was generated, the first using METHOD = COMPLETE (E4) and the other using the default option (E5). The summary of the results are given in (E6). In both cases, the ANOVA in the QUICK CLUSTER option showed significant difference between cluster center means for all variables (E7). A discussion of the respective results is next. If appropriate, two levels of clustering will be identified.

Discussion

A] Cluster using COMPLETE

Six clusters were identified at a lower level, revealing the commonalities among the member institution programs (E6). The first cluster had 1986 as a median date of establishment. The range of years was 1976-1993. It is characterized primarily by Engineering Management Programs based in Engineering Schools who offer mostly masters degrees. Cluster 2 has 1991 as a median date of establishment, with range 1981-1994 offering mostly masters degrees in Management of Technology. These are either business school based, or are joint programs in business and engineering schools. The third cluster has a median date of establishment of 1982 with range 1975-1984. These are mostly Engineering Management Programs offering masters degrees and are jointly housed by business and engineering schools. Cluster 4 is similar to Cluster 1, except that their median date of establishment was 1975, as compared to 1986 in the latter. Cluster 5 consisted of the oldest programs which were mostly neither called Engineering Management nor Management of Technology, and offered relatively more bachelors degrees than any other cluster. The last cluster consisted of older programs coined Engineering Management, are housed in engineering schools, and offered mostly masters degrees.

Clustering into 3 groups show that Clusters 1 and 2 can be joined, with proximity of dates of establishment as the link if the dendogram is followed (E4). It could also be argued that Clusters 1 and 4 can be joined based on the Engineering Management label and the academic base. Clusters 3 and 4 (E4) can be joined into a cluster

using their proximity based on date and program title. Finally, the fifth and sixth clusters can be grouped together with date and academic base as the similarity base.

B] Cluster using BAVERAGE (default)

Clustering using this method is more difficult as there are more blurred distinctions among the groups. However, 6 clusters can also be specified. This allows comparison with the other method.

Cluster 1 can be defined as Engineering Management Programs housed in engineering schools, offering masters degrees, and having 1984 as their median date of establishment, with range 1976-1993. This cluster had the most number of members. Cluster 2 can best be described as Management of Technology programs which are newly established in business schools. They offer mostly masters degrees. The third group consists mainly of joint programs, and the program titles are equally divided between Engineering Management and management of Technology. The distinct characteristic of cluster 4 is that they offer more bachelors degrees than any other cluster. Cluster 5 is made up of older Engineering Management Programs in engineering schools offering masters degrees. The last cluster were the oldest programs, having no Engineering Management nor Management of Technology titles.

Reducing the number of clusters will join the second and third clusters based on the proximity of their dates of establishment. Consequently, clusters 5 and 6 can be joined with the same reasoning. Clusters 1 and 5 only differ by their years of establishment. While Cluster 3 is distinguished by its joint programs, cluster 4 by the number of bachelors degrees they offer.

C] Analysis

The degrees offered was a factor in only one instance, Cluster 4 of the default method, where the number of bachelors degrees was the predominant characteristic. Most ETMs offer masters degrees, and in most clusters, this was not the difference. In both methods, two groups of Engineering Management programs based in engineering schools were separated by their dates of establishment. Both methods identified a cluster by the predominant occurrence of joint programs. However, the COMPLETE method mostly included Engineering Management programs while the BAVERAGE method identified as many Management of Technology programs. Clusters 5 and 6 in both methods were fairly consistent, characterizing the 2 oldest groups. Overall, both methods gave fairly consistent results. The clusters were more evident in the COMPLETE method when viewing the dendograms. However, the 'degrees offered' variables hardly entered the clustering decision. The BAVERAGE method clustered a group using the BS offered. Either methods can be used, depending on what the decision maker is looking for, and how the results can be used.

D] Extensions of the Study

This analysis is just an initial attempt to classify ETMs according to their broad characteristics. Other variations of cluster analysis can be performed to check for other ways of classifying ETMs. Different distance measures may be used in combination with other cluster methods. Other variables in the original survey instrument may be included in lieu of the degrees offered (BS, MS, PHD). They may also be added to the full model used.

Conclusion

The cluster analysis conducted show that ETM programs can be classified using date of establishment, title of program, academic base, and to a lesser extent, type of degrees offered. The classification schemes were not predominated by any of the variables. Instead, a combination of the variables worked well to determine the classifications. The BAVERAGE and COMPLETE methods, in general, give similar results.

REFERENCES

1. Hair J., Anderson R., Tatham R., & Black W., MULTIVARIATE DATA ANALYSIS with Readings, New York: Macmillan Pub. Co., 1992.

2. Norusis M., SPSS/PC+ Base System User's Guide, Version 5.0, Chicago: SPSS INC. ,1992.

3. Norusis M., SPSS/PC+ Professional Statistics, Version 5.0, Chicago: SPSS Inc., 1992.