Naslov članka/Article:

A Study on Clustering High Schools According to Multiple Success Variables

Študija povezovanja srednjih šol na osnovi raznovrstnih, z uspehom povezanih spremenljivk

Avtor/Author: Sultan Uçar-Altun and Mualla Aksu

DOI https://doi.org/10.59132/vviz/2019/3/3-24

Vodenje v vzgoji in izobraževanju 44, 3/2019, letnik 17
ISSN 1581-8225 (tiskana izdaja)
ISSN 2630-421x (spletna izdaja)

Izdal in založil: Zavod Republike Slovenije za šolstvo
Kraj in leto izdaje: Ljubljana, 2019

Spletna stran revije:
https://www.zrss.si/strokovne-revije/vodenje-v-vzgoji-in-izobrazevanju/

A Study on Clustering High Schools According to Multiple Success Variables

Sultan Uçar-Altun
Korkuteli Middle School, Turkey

Mualla Aksu
Akdeniz University, Turkey

This paper presents the results of a relational, descriptive study. Its main purpose is to examine variables that determine multiple success categories in high schools. The sample consists of 80 public high schools in Antalya. Data were collected through the school information form developed by the researchers, based on the related literature and data available at school. Academic Score (AS), Social, Cultural, and Artistic Score (SCAS), and Sportive Score (SS) were used as success categories. Four clusters were obtained by cluster analysis. Clusters were called Academically Powerful Schools (APS), Schools in Need of Improvement (SNI), Sportively Powerful Schools (SPS), and Socially, Culturally, and Artistically Powerful Schools (SCAPS). According to the results, APS was the best in AS, SNI was the worst except AS, SPS was the best only in SS, and SCAPS was the best only in SCAS. Science high schools were clustered in APS. Anatolian high schools were located in four different clusters, half of them in SNI. Vocational and technical Anatolian high schools were located in three different clusters, more than half of them in SNI, but none in APS. Anatolian imam and preacher high schools were located in two different clusters. Only one-third of them were clustered in SCAPS while two-thirds were in SNI. Multi-program Anatolian high schools were also clustered in SNI only. In the study, significant correlation coefficients were obtained among some selected success variables.

Keywords: multiple success variable, high school, cluster analysis, correlation coefficient

Introduction and Theoretical Framework

Manpower is the most valuable resource of a country. Schools are institutions where these valuable resources are turned into power. Today, individuals with scientific thinking skills, productivity, creativity and the ability to solve the problems faced are very much needed. Schools have social, political and economic function. The social function of a school is to socialize the individuals and to de-

velop the culture. The political function is to educate youth to be loyal citizens while the economic function is to meet the needs of the economy in terms of not only manpower and but also brain power.

The Basic Law of National Education No. 1739 issued in 1973 determines general framework of Turkish national education system. According to this law, the overall aim of the education system is as follows: (1) to promote the welfare and happiness of the citizens, (2) to support and accelerate economic, cultural and social development in national unity and cohesion, and (3) to make the nation a constructive, creative and distinguished partner of contemporary civilization (MEB 2005). Education system in Turkey consists of two main parts: formal and non-formal education. Formal education is the regular schooling conducted within schools for individuals in a certain age group and includes pre-primary, primary, secondary and higher education institutions. In this context, educational institutions at all levels have their own specific objectives. 'The aim of secondary education is to give students a minimum level of common culture, to acquaint them with problems of the individual and society, to teach them how to seek solutions, to raise awareness in order to ensure their contribution to the socio-economic and cultural development of the country and to prepare students for higher education, for professions, for life and for business in line with their interests and skills' (MEB 2005).

Secondary education institutions, which are the subject of this study, are called high schools. Various programs are applied in high school education and high schools are given different names based on their program diversity. High schools in the Turkish education system are the following: Science High School (SHS), Anatolian High School (AHS), Vocational and Technical Anatolian High School (VTAHS), Anatolian Imam and Preacher High School (AIPHS), Multi-Program Anatolian High School (MPAHS), social sciences high school, fine arts high school and sports high school.

SHS provides a basis for upbringing students as scientists in the field of science and mathematics. Similarly, social sciences high school provides a basis for upbringing students as scientists in the field of social sciences. On the other hand, AIPHS aims at providing the necessary knowledge and skills that will be the source of religious services such as imamate, preaching and teaching in Qur'an courses. VTAHS aims at raising the labor force needed by the labor market in the fields of industry, trade, textile, construction, tour-

ism, chemistry, agriculture, and health. Youth in those schools is also being prepared for employment by providing them the possibility to gain the spirit of entrepreneurship, professional ethics, occupational health and safety, social and environmental responsibility, and work habits (MEB 2013).

AHS is open to enable students to prepare for higher education programs according to their talents and achievements and to learn foreign languages at a level that enables them to follow scientific and technological development in the world (MEB 1999). Multiprogram high school is open in accordance with objectives, principles and policies of National Education Basic Law, development plans, government programs and decisions of national education council, by taking into consideration education needs, student potential, and education cost of small settlements (MEB 2001).

At this point, the following question becomes important: to what extent are these schools able to achieve their goals and to what extent do the same types of schools provide expected outcomes? Evaluating the outcomes of the curriculum and using the information obtained as a result of these evaluations in the program development are of great importance for effectiveness of the education system (ERG 2019). The Ministry of National Education (MONE) has a huge quantity of data on the whole school system and educational data mining may be benefited in order to get the information needed. García et al. (2011) define the term 'educational data mining' as the process of converting raw data obtained from the educational systems into information that can be used by educational software, program developers, educational administrators, decision makers, teachers and researchers. Educational data mining is a new discipline that develops methods to examine increasingly large-scale data of the original type from educational organizations and uses them to better understand students and educational organizations (see http://www.educationaldatamining.org). It can be used to provide managers with the data-based information they need to increase the effectiveness and efficiency of educational organizations. With this information, it is possible to analyze the data, increase student achievement by revealing the reasons for student success and failure, identify problems in educational environments, and create more effective environments (Özbay 2015). Student learning data are being explored to develop predictive models by applying educational data mining methods. These models play an important role in developing adaptive learning systems. This way, adaptations or interventions based on

the model predictions can be used to modify student experience next or to suggest additional academic services to support learning (Bienkowski, Feng, and Means 2012). Especially in the field of education, there is a huge quantity of data about students, teachers, teaching environments, measurement and evaluation results. Such reality clearly shows the importance of using the information stored in these data collections in order to explore patterns and improve the efficiency and quality of education.

In education, data mining techniques, such as classification and clustering, are usually used to categorize students, based on the kinds of personal learning data, on student demographic data, or both (Bienkowski, Feng, and Means 2012). Cluster analysis is a multivariate statistical technique the main purpose of which is to group objects according to their characteristics (Hair et al. 2010). 'Clustering is the process of examining the properties of objects and grouping them into clusters according to some distance measurements. The aim of the cluster analysis is to collect objects close to each other in the same cluster, while collecting distant objects in different clusters' (Leskovec, Rajaraman, and Ullman 2014, 241). Likewise, James et al. (2013) described clustering analysis as the process of dividing objects into different groups, with observations within each group being very similar. Thus, while similar data enters the same cluster, different data is contained in different clusters (Singh and Singh 2012). Clustering examines a collection of points, and groups the points into clusters according to some distance measure. This means that points in the same cluster are spaced at a small distance from the others, while points in different clusters are spaced at a large distance (Leskovec, Rajaraman, and Ullman 2014). However, it needs to be revealed what it means for two or more objects to be similar or different in order to make this procedure concrete (James et al. 2013).

A limited number of studies on cluster analysis have been found in the literature. Among these researches, those related to education are summarized below: Perry (2000) conducted a research aiming at aggregating and summarizing data from Virginia public school districts, creating a paradigm that will quantify and rank order the variables, and place school districts into groupings. In the study, schools were clustered into two groups using the k-means cluster analysis procedure. According to the results, Virginia is experiencing shortages of instructional personnel, especially in the field of special education, mathematics, science, and technology endorsement. The most significant variables were the competition

from other school districts, retirement, efforts to reduce teacher to pupil ratios, and salaries. Green (2017) aimed at developing a typology of teachers and identifying the groups of teachers that perform better in regard to student achievement. In the study, three teacher profiles, high, mid, and low were found. The high cluster typically consisted of younger teachers, and predominately male. The older, with the majority of female educators were placed in the mid cluster while educators towards the end of their careers were found in the low cluster. On the other hand, Ungricht (1997) identified the relationship between learning strategies and demographic and educational performance variables, and explored patterns of learning of distinct clusters. The result shows that distinct groups exist among learners and that learning strategies are linked to specific educational performance measures. Conducting a study in Maryland, New Jersey, Pennsylvania, and Virginia, Warren (2007) used hierarchical clustering analysis and created five clusters of school divisions. The findings of the study indicate that there is a relationship between incidents related to student discipline and student achievement. Similarly, Halsell (2007) examined the performance of schools within homogeneous clusters by considering that a school's dominant student population will have a significant influence on academic performance. Schools were classified on the basis of dominant student populations and determinations were made concerning statistically significant differences in mean reading and math scores. According to the results, schools did demonstrate significant differences in reading and math scores with selected schools performing significantly above expectations, certain schools performing significantly below expectations, and many demonstrating no significant difference relative to similar populations located in homogeneous clusters.

Studies conducted by applying clustering analysis in various fields can produce different results based on preferred variables for the statistical usage. For example, Page (2004) investigated the impact of charter schools on student achievement by using cluster analysis in North Carolina and found that a four-cluster typology was the most 'optimal' for the study although there was a problem in internal validity. Page's primary concern was related to the lack of similarity between the charter and non-charter schools within the four clusters. Humphreys (2006) explored the institutional characteristics of California charter schools and the extent to which these characteristics relate to student achievement. Findings demonstrate that charter schools differ in both

teaching and student characteristics. Three clusters were found named 'under-resourced,' 'affluent,' and 'specialty' schools. Surprisingly, the under-resourced cluster had the highest academic performance on four different measures of growth. Teacher's length of service negatively predicted student achievement while student/teacher ratio positively predicted achievement scores. Crain-Dorough (2003) conducted a three-phase study in order to examine the characteristics of student dropouts and the characteristics of schools successful and unsuccessful in mediating dropouts. Three clusters of schools were found, named 'high achievers,' 'average achievers,' and 'low achievers.' As for the dropouts, three clusters were found, namely 'quiet dropouts,' 'typical dropouts,' and 'high-achieving pushouts.' Significant differences were found among the set of dependent variables such as attendance rate, class size, student achievement, suspension rate, teacher certification, and teacher test scores. The results showed that consistently low dropout schools had significantly higher student achievement than the less effective schools, while the more effective schools had significantly higher attendance rates and student achievement than the consistently high dropouts schools.

In the context of e-state in Turkey, big and various data collections are related to teachers and students and used by MONE; E-school, Educational Information Network, and MONE Information Systems. Unfortunately, these data were not used beyond reporting and they were not benefited sufficiently in order to obtain the information needed. However, MONE initiated a new study and focused on this issue in its 2023 Education Vision Document published in 2018. In this document, the emphasis is put on processing data obtained from all the levels of education and on using them in the process of decision-making, planning and evaluation. In order to achieve these targets, the following tasks were determined (MEB 2018):

- Initiation of 'Data Based Planning and Management System' at the school level for monitoring, evaluation and development of management and learning activities throughout the country.
- Integration of data from the Ministry's current systems into an easily accessible Educational Data Warehouse.
- Establishing an online platform where the Ministry and school administrators can monitor school development plans across county, province, region and country.

- Establishing a Geographic Information System to determine the capacity of schools when planning educational resources.
- Justice-based allocation of resources provided by the Ministry.
- Establishing a decision support mechanism to determine which school will be supported and what kind of support will be provided by following the school profile evaluation data.

All this knowledge reveals the need for identifying the levels of multiple success in high schools that are expected to prepare students for higher education or life experience. In general, there are very few studies investigating the impact of public schools on student achievement (Page 2004). However, many variables, which may be controlled or not, have an impact on student success; they are caused by students themselves and outside themselves (Demirtaş 2010). In light of this information, a study was designed to determine how high schools are clustered according to their multiple success levels and to compare schools in different clusters in terms of some variables.

General purpose of this study is to cluster high schools according to their multiple success levels and to compare them in terms of some variables. To achieve this purpose, the following research questions were identified:

1. How are high schools clustered according to multiple success variables?
2. Which type of high schools is situated in which clusters?
3. How are clusters compared in terms of some variables?
4. Are there significant correlations between selected two variables in terms of certain demographic characteristics of high schools?

Methodology

This is a relational, descriptive study. The main purpose of a descriptive study is to identify the state of a situation or phenomenon and explore the relationships between phenomena (Christensen, Johnson, and Turner 2015). In this context, variables that determine multiple success categories in high schools were examined in the study.

The population consists of 202 public high schools in Antalya. Social sciences, sports and fine arts high schools were not included

TABLE 1 Descriptive Data on the Population and the Sample

School Types	Population		Sample	
	N	%	*n*	%
Science High School	11	5.56	5	6.25
Anatolian High School	76	38.38	30	37.50
Vocational and Technical Anatolian High School	63	31.82	28	35.00
Anatolian Imam and Preacher High School	36	18.18	12	15.00
Multi-Program Anatolian High School	12	6.06	5	6.25
Total	198	100.00	80	100.00

in the field of inquiry due to their small number and differentiation from others in terms of admission requirements. Data were collected through the School Information Form (SIF) developed by the researchers based on the related literature and data available at schools. Since four schools were combined and closed, SIF were sent to 198 high schools. After incomplete incorrectly filled forms were excluded, analyses were made from the data of 80 schools. Table 1 illustrates descriptive data on the population and the sample by school types. As shown in the table, there are no big differences between the population ratio and the sample ratio by school types.

Rapid Miner for cluster analyses and SPSS 23.00 for correlation coefficients were used in order to analyze data. The variables of Academic Score (AS) used in cluster analysis are school mean on university entrance exam, school ratio of achievement certificates, and school ratio of appreciation certificates. The variables of Social, Cultural, and Artistic Score (SCAS) used in cluster analysis are school ratings obtained from the contests on poetry, essay, story, visual arts, painting, and sculpture as well as individual and team music competitions. The variables of Sportive Score (SS) used in cluster analysis are the results obtained individually and in team sports such as football, volleyball, basketball, tennis, table tennis, weightlifting, wrestling, and swimming.

Results

In this sub-title, data on the research questions were presented in tables and the results obtained in the study were explained.

Clustering High Schools According to Multiple Success Variables

In this study, accepted criteria were implemented in order to determine success scores used for clustering. In this context, three

TABLE 2 Quotient Criteria for District, Provincial, and Country Level Awards

Awards	District	Province	Country
Sport	1	4	8
Painting	2	4	10
Music	2	4	10
Literature	2	4	10
Projects	–	4	10

types of scores named Academic Score (AS), Social, Cultural, and Artistic Score (SCAS), and Sportive Score (SS) were used as success categories. The formula used for AS is as follows:

$$\begin{aligned} \text{AS} = &\text{ normalized score of achievement and appreciation} \\ &+ \text{normalized point average of the school.} \end{aligned}$$

Students get certificate of achievement when they obtain mean scores among 70–84. The formula of (70 + 84)/2 = 77 was used to obtain the certificate of achievement score. The score obtained from this procedure was multiplied by the number of certificates given at school. Then the obtained value was divided by the total number of students at school.

Students get certificate of appreciation when they obtain mean scores among 85–100. The formula of (85 + 100)/2 = 92.5 was used to obtain the certificate of achievement score. The score obtained from this procedure was multiplied by the number of certificates given at school. Then the obtained value was divided by the total number of students at school.

Quotients were determined for SCAS and SS based on characteristics of the awards gained by individual students or student teams at school. Expert opinion was obtained to determine quotient criteria shown in table 2. The scores obtained were normalized in the 0–1 range when determining clusters.

K-Means clustering algorithm was used for data analysis. Eighty schools were subjected to statistical analysis for clustering according to their scores in three categories. Cluster analyses were made five times by using the number of clusters from three to seven. As a result of cluster analysis, schools were classified into four clusters (numbered 0–3) with expert opinion. Table 3 illustrates Euclidean distance in cluster analysis. Cluster 0 consists of 10 schools. In this cluster, all scores obtained for AS, SCAS and SS are larger than average at a ratio of 192.43 percent, 54.04 percent, and 52.01 percent, respectively. Cluster 1 consists of 45 schools. In this cluster,

TABLE 3 Euclidean Distance in Cluster Analysis

Cluster	Distance	Average dist.	Notes
0	10	0.155	AS is on average 192.43% larger SCAS is on average 54.04% larger SS is on average 52.01% larger
1	45	0.023	SCAS is on average 64.00% smaller SS is on average 40.67% smaller AS is on average 28.94% smaller
2	9	0.072	SS is on average 188.00% larger AS is on average 33.10% smaller SCAS is on average 0.62% smaller
3	16	0.066	SCAS is on average 148.45% larger SS is on average 23.87% smaller AS is on average 20.27% smaller

TABLE 4 Cluster Model

Cluster	AS	SCAS	SS
Academically powerful schools (APS)	0.840	0.344	0.264
Schools in need of improvement (SNI)	0.204	0.079	0.103
Sportively powerful schools (SPS)	0.192	0.222	0.501
Socially, cult., and artistic. powerful schools (SCAPS)	0.229	0.556	0.132

all scores obtained for AS, SCAS and SS are smaller than average at a ratio of 64.00 percent, 40.67 percent, and 28.94 percent, respectively. Cluster 2 consists of 9 schools. In this cluster, only SS is larger than average at a ratio of 188.00 percent while AS and SCAS are smaller than average at a ratio of 33.10 percent and 0.62 percent, respectively. Cluster 3 consists of 16 schools. In this cluster, only SCAS is larger than average at a ratio of 148.45 percent while SS and AS are smaller than average at a ratio of 23.87 percent and 20.27 percent, respectively.

Four clusters obtained by cluster analysis are given names according to their success differences within the three categories. Table 4 illustrates the names of the cluster as well as the mean score of schools in each cluster in the range of 0–1. Cluster 0 was called Academically Powerful Schools (APS), and APS was found to have the best AS. Cluster 1 was called Schools in Need of Improvement (SNI), and SNI had the worst scores except AS. Cluster 2 was called Sportively Powerful Schools (SPS), SPS had the best score only in SS. Cluster 3 was called Socially, Culturally, and Artistically Powerful Schools (SCAPS), and SCAPS had the best score in SCAS only.

TABLE 5 Numbers of Members in Four Clusters by School Types

Clusters	School Types					
	AHS	AIPHS	MPAHS	SHS	VTAHS	Total
1. Academically Powerful Schools (APS)	5	–	–	5	–	10
2. Schools in Need of Improvement (SNI)	15	8	5		17	45
3. Sportively Powerful Schools (SPS)	3	–	–		6	9
4. Socially, Culturally, and Artistically Powerful Schools (SCAPS)	7	4	–		5	16
Total	30	12	5	5	28	80

Distribution of High Schools into Different Clusters According to Their Types

As illustrated in table 5, all five science high schools included in the study were clustered in APS. Anatolian high schools were located in four different clusters and half of them were placed in SNI. Vocational and technical Anatolian high schools were located in three different clusters and 17 out of 28 schools were placed in SNI while no school was in APS.

Anatolian imam and preacher high schools were located in two different clusters. Only one-third of them were placed in SCAPS while two-thirds were in SNI. Multi-program Anatolian high schools were also clustered in SNI only.

Comparing Different Clusters in Terms of Some Variables

In order to compare different clusters in terms of some variables, descriptive data on 16 variables were examined. As illustrated in table 6, APS was found to be the best for nine variables. However, APS was the worst for variables 'information technology,' 'number of students per teacher,' and 'number of students per classroom.' On the other hand, SNI was the best only for two variables 'laboratory facilities' and 'sport facilities.' SPS was the best for variables 'ratio of white flag,' 'information technology,' and 'teachers' length of service at school.' However, it was the worst for variables 'laboratory facilities,' 'sport facilities,' 'ratio of teachers with graduate education,' and 'ratio of certificates of achievement and appreciation.' SCAPS was only the best for two variables, namely 'number of students per teacher' and 'number of students per classroom.' Nevertheless, it was the worst for variables 'ratio of white flag,' 'average teachers' length of service,' and 'teachers' length of service at school.'

TABLE 6 Comparisons on the Characteristics of Four Clusters According to the SIF Variables

Items	Clusters			
	(1)	(2)	(3)	(4)
Ratio of white flag (%)	80.00	73.33	88.90	68.80
Laboratory facilities	11.2	12.9	7.0	11.97
Having information technology	1.2	6.1	7.5	3.2
Music, painting classroom	6.5	2.1	2.5	6.2
Sport facilities	10.2	12.1	9.1	9.4
Number of students per classroom	28.5	22.1	23.4	20.0
Ratio of hourly paid teachers (%)	0.0	8.8	8.4	6.7
Ratio of teachers with graduate education (%)	9.4	9.1	7.0	9.0
School's admission score	452	235	237	274
School's mean on university entrance	357	196	203	206
Ratio of certificates of achievement and appreciation	87	25	22	27
Ratio of disciplinary punishment (%)	0.0	6.8	4.4	3.2
Days of absence without excuse	4.90	6.20	5.78	5.19
Average teacher's length of service	19.2	15.2	16.2	12.3
Teachers' length of service at school	6.40	6.62	8.80	4.60
Number of students per teacher	14.5	12.1	12.1	11.4

NOTES Column headings are as follows: (1) academically powerful schools, (2) schools in need of improvement, (3) sportively powerful schools, (4) socially, culturally, artistically powerful schools.

Correlations between Two Selected Variables in Terms of Some Demographic Characteristics of High Schools

Correlation coefficients were calculated in order to determine relationships between some selected demographic characteristics of high schools. These values were also interpreted based on the criteria recommended by Cohen, Manion, and Morrison (2007) for the effect size of correlation coefficients.

As illustrated in table 7, sixteen correlation coefficients were calculated. In order to determine the common variance between the two variables, it was necessary to calculate the effect size and for this purpose, the value obtained from the square of the correlation coefficient was used. The square of the correlation coefficient shows the proportion of variance in one variable that can be attributed to its linear relationship with the second variable. In other words, it indicates the amount the two variables have in common (Cohen, Manion, and Morrison 2007).

Although all correlation coefficients were significant at $\alpha = 0.05$ level, eleven of them had weak effect sizes. A strong positive relationship ($r^2 = 0.57$) was found between variables 'academic score' and 'school admission score.' It means that 57 percent of the vari-

TABLE 7 Comparisons on the Characteristics of Four Clusters According to the SIF Variables

Variable 1	Variable 2	r	P	r^2	Int.*
Academic Score	School's admission score	0.75	0.000	0.56	Strong
	Ratio of hourly paid teachers	–0.45	0.000	0.20	Modest
	Average teacher's length of service	0.29	0.008	0.08	Weak
	Ratio of disciplinary punishment	–0.23	0.039	0.05	Weak
	Music, painting classrooms	0.24	0.027	0.06	Weak
Social, Cultural, & Artistic Score	Music, painting classrooms	0.25	0.023	0.06	Weak
	School's admission score	0.22	0.042	0.04	Weak
	Ratio of disciplinary punishment	–0.23	0.040	0.05	Weak
Sportive Score	Average teacher's length of service	0.42	0.000	0.17	Weak
	Teacher's length of service at school	0.26	0.016	0.07	Weak
	Ratio of hourly paid teachers	–0.24	0.027	0.06	Weak
Ratio of certificates of achievement and appreciation	School's mean on university entrance exam	0.76	0.000	0.57	Strong
	School's admission score	0.72	0.000	0.52	Strong
School's mean on university entrance exam	School's admission score	0.69	0.000	0.47	Moderate
Days of absence without excuse	School's admission score	–0.26	0.020	0.07	Weak
	Ratio of certificates of achievement and appreciation	–0.23	0.042	0.05	Weak

ation shown by academic score can be attributed to the tendency of academic score to vary linearly with school admission score. 'Academic score' had a modest negative correlation (r^2 = 0.20) with variable 'ratio of hourly paid teachers' but a weak negative correlation (r^2 = 0.05) with variable 'ratio of disciplinary punishment.' Other strong effect sizes were calculated in correlations between variables 'school admission score' and 'ratio of certificates of achievement and appreciation' (r = 0.72) as well as between variables 'ratio of certificates of achievement and appreciation' and 'school mean on university entrance exam' (r = 0.76). A modest effect size was found in the negative correlation (r = –0.45) between variable 'ratio of hourly paid teachers' and 'school admission score.' Not only sportive score but also social, cultural, and artistic score had weak relationships with the selected variables, and their effect sizes ranged from 0.04 to 0.17 (table 7).

As shown in table 7, 'school mean on university entrance exam' can explain 57 percent of the variation shown by 'ratio of certificates of achievement and appreciation' with a strong effect size. Similarly, school admission score can explain 52 percent of the variation shown by 'ratio of certificates of achievement and appre-

ciation.' In addition, 47 percent of the variation of 'school mean on university entrance exam' can be explained by 'school admission score.' On the other hand, weak correlation coefficients were obtained between the variable 'days of absence without excuse' with both variables 'school admission score' (r = –0.26) and 'ratio of certificates of achievement and appreciation' (r = –0.23).

Discussion and Conclusion

In this study, public high schools in Antalya province were examined in terms of clusters according to multiple success variables based on 2018 data. The aim of the study was to determine success clusters of schools and to compare schools in different clusters in terms of some selected variables. Clustering analysis was applied by using K-Means algorithm with AS, SCAS, SS obtained from the data collected from schools. At the end of the analysis, it was determined that schools were divided into four clusters called APS, SNI, SPS, and SCAPS.

The APS cluster consisted of 10 schools that were above average by all scores. It means that only 11.25 per cent of schools in Antalya were found successful in all three categories. On the other hand, the SNI cluster consisted of 45 schools that were below average by all scores. The fact that the SNI cluster, consisting of 56 percent of 80 schools included in the study, had scores below the average in all success categories, shows the real dimensions of differences between schools in Antalya. Turkey has a quality problem in education in terms of standards accepted on international levels. In addition, success in education is not disseminated homogeneously across the country. In other words, there are schools with success levels that range from the highest to the lowest in the same region, in the same province and even in the same district (Önder 2012). It is known that the MONE authorities also see this issue as an important problem. Quality differences among schools in Turkey are a fundamental problem, and the differences between good and poor schools are so large that this fact can be called a quality gap. However, it is one of the tasks of the state to provide equal opportunities for all children to receive the same level of education (Yıldız 2015). Therefore, taking measures to eliminate quality and quantity differences among schools is a necessity, and the General Directorate of Secondary Education in MONE is responsible for this task. Quality differences between the types of schools in which students are selected via central examination and the open

admission schools are quite high. This fact indicates that opportunities are not evenly distributed. Such a situation contradicts the principle of equal distribution of opportunities in education (Berberoğlu and Kalender 2005).

Kurebayashi (2015), one of faculty members at Tokoha University, visited some schools within the scope of a project in Turkey, and determined by observation that there were big differences among the quality of schools in various regions. According to Kurebayashi (2015), these differences should first be eliminated and qualified teacher training should be addressed. By drawing attention to this problem in the 2023 Education Vision Document, MONE (MEB 2018) suggested a student achievement follow-up research in order to reduce the differences between schools and regions and to see the education system as a whole. In this document, the need for setting up a 'Geographic Information System' was also expressed in order to determine capacities of schools in terms of educational resources.

In this study, all science high schools were clustered in APS. This result shows that science high schools are the most successful schools in terms of achieving their goals. Although some Anatolian high schools were placed in the best cluster together with science high schools, they were also distributed to other three clusters. The reason of this may be that all general high schools were converted into Anatolian high schools after 2013. Two thirds of Anatolian imam and preacher high schools were clustered in SNI, and only one third were in SCAPS cluster. No Anatolian imam and preacher high schools were in APS and SPS clusters. Similarly, six out of ten of vocational and technical Anatolian high schools included in the study were in SNI cluster. There were no vocational and technical Anatolian high schools in APS cluster. All of five multi-program Anatolian high schools were also classified in SNI cluster. Likewise, Bilen et al. (2014) conducted a cluster research by using data of the students who took the university entrance examination in Istanbul. As a result of their study, it was determined that schools were grouped into five clusters in terms of student achievements at the university entrance examination. Science high schools in their study were classified in the best cluster, and the success ranking continued by Anatolian high school, general high school, Anatolian imam and preacher high school and vocational and technical Anatolian high school. In the study conducted by Köse (1999), schools were ranked according to university entrance scores and based on the university registration rates,

vocational and technical high schools were found the least successful. By applying K-means clustering algorithm on the training data, Shovon, Islam, and Haque (2012) grouped students in three classes; high, medium and low, according to their grade. As could be seen in the above-mentioned studies, schools were classified using only few variables in terms of student achievement. In this context, the current study based on multiple variables data differs from the previous ones.

In this study, APS cluster is seen to be disadvantageous in terms of the number of students per teacher and the number of students per classroom although school admission scores and school mean at university entrance exam in APS cluster are higher than the others. Nevertheless, the majority of students and their parents prefer high schools in this cluster and student quota for each classroom gets full all the time.

When compared with other clusters, in APS both disciplinary and absenteeism rates are the lowest while the rate of achievement and appreciation certificates is the highest. This result is consistent with the following findings of the research conducted by MONE and UNICEF (MEB 2011, 5): 'The presence of children in school ensures that they are protected against all kinds of risks that may come from outside and that they can obtain educational attainment from school. The fact that children stay away from school affects all aspects of their development negatively, and makes it difficult for them to meet with appropriate prevention and intervention programs timely.' However, absenteeism, grade retention and dropout, which negatively affect one's participation in education and the quality of education they receive, are very common in secondary education (ERG 2011). Student absenteeism is a growing problem in public schools and the adverse effects have been well established through research (Grant 2016). Chronically absent students are more likely to experience negative academic outcomes, such as grade retention and dropping out, compared to their peers who consistently attend school (Wallace 2017). According to Jermain (2018), absenteeism not only affects the student who does not attend school but also affects those around the student. Chronic absenteeism is an indicator something is wrong and should be addressed as soon as possible. Commitment to school, on the other hand, affects student attendance and academic performance (Boesel 2001). For this very reason, schools with low suspension rates report using a variety of successful incentive programs to promote positive academic and social behavior (Christie,

Nelson, and Jolivette 2004). It can be seen that the results of this current study are consistent with the above-mentioned literature since APS has the lowest and SNI has the highest absenteeism rate.

In another comparison, it was seen that the school admission scores and school mean on university entrance exam were lower in SNI cluster than the others. Whereas SNI cluster had more sports and laboratory facilities. SCAPS cluster was only advantageous in terms of the number of students per classroom and the number of students per teacher. Surprisingly, SPS cluster was found to be disadvantageous in terms of sports facilities. It was also disadvantageous in terms of laboratory facilities, ratio of teachers with graduate education, and achievement and appreciation score while it was advantageous in terms of white flag and teachers' length of service at school. Although two-thirds of Anatolian imam and preacher high schools were clustered in SNI, they were in the best position in terms of 'number of students per classroom,' and 'number of students per teacher' variables. Unfortunately, science high schools, all clustered in APS, were found in the worst situation in terms of the same variables. The fact remains that although the success level of such schools is still low, MONE has been allocating much more resources to religious education for years. In 2017 budget proposal, allowance per student for imam and preacher high schools was twice as high as for general secondary schools (ERG 2017). However, not only teachers and principals but also parents and community expect a fair budget and a balanced support by MONE for all types of schools.

As it is well known, the teacher has a vital role in education. In this study, APS is found to have the highest number of experienced teacher, being also paid the lowest. On the other hand, SNI has the highest number of paid teachers and SCAPS has the lowest number of experienced teachers. Similarly, in South African education, there are gross differences noticeably in the per capita spending, the availability of physical facilities, the average qualifications of teachers and the pupil/teacher ratios (Turnage 2011). Stiefel, Schwartz, and Iatarola (2001) report that low performing schools have higher teacher/pupil ratios, employ teachers who are proportionately less licensed, less paid, and less experienced. In Turkey, the rotation of teachers among schools, especially in disadvantaged areas is higher because it is difficult to keep qualified employees working in hard conditions for a long time. Unfortunately, MONE prefers to employ paid teachers instead of experienced ones, to cover the numerical deficit (Önder 2012).

Since this study was carried out in 80 high schools in Antalya, it has some limitations and no doubt the results cannot be generalized to the entire Turkish education system. However, the number of clusters can be re-examined in similar studies conducted in the future. At the same time, MONE may benefit from this study in evaluating the performance of the schools. As targeted in 2023 Education Vision Document (MEB 2018), school needs should be determined based on the data such as general and special classrooms, gymnasium, numbers and qualifications of teachers. In this way, ensuring a fair distribution of resources can reduce differences between schools. Based on student population and preference system, undesired schools can be transformed into different types of schools. Schools implementing vocational programs can be organized by taking into account the needs of industry and society employment area in the country. Similarly, Perry (2000) asserts that practical application of cluster identification is to enable educational agencies to develop targeted intervention strategies that address common issues within clusters. In this way, targeted and focused intervention strategies would facilitate the most efficient allocation of resources.

Creating elite secondary education institutions and considering the fact that these institutions are seen as a means of going to university, causes serious inequalities in Turkish education system, especially in practice and social perception. Indeed, providing elitist education to a small number of students in a small number of schools poses a major threat to social cohesion and economic development. This situation is no longer sustainable and threatens the future of all segments of the society as a whole. Scores obtained at the national scale are generally evaluated in terms of only ranking and the meaning of knowledge and skills levels for individuals and the society is not emphasized. Unfortunately, evaluation shows that only 2 per cent of students across Turkey could be given 'good education' (TED 2008). Success differences between schools could be reduced by giving every child access to better education in a more egalitarian system. For this reason, it is necessary to allocate government resources based on school conditions instead of providing general equality in the distribution of public resources. Not following the school success equals waste of resources. In this respect, a structure that will systematically monitor student success and ensure measures when the school falls below the standards is also needed (Önder and Güçlü 2014).

As a conclusion, it is expected that education authorities on both provincial and central levels may benefit from the findings of this study in order to discover the actual school conditions and to improve the unsatisfactory areas. The study can also contribute to more effective and need-oriented usage of human resources and money allocated to education. However, considering that the current study does not even include all high schools in Antalya province, it can be clearly seen that further research is needed in this subject.

Acknowledgments

This paper was produced from a part of PhD dissertation of the first author and conducted under supervision of the second author. The study was supported by Akdeniz University with SDK-2018-4223 project code.

References

Berberoğlu, G., and İ. Kalender. 2005. 'Öğrenci başarısının yıllara, okul türlerine, bölgelere göre incelenmesi: ÖSS ve PISA analizi.' *Eğitim Bilimleri ve Uygulama* 4 (7): 21–35.

Bienkowski, M., M. Feng, and B. Means. 2012. *Enhancing Teaching and Learning through Educational Data Mining and Learning Analytics.* Issue in Brief 1. Washington, DC: US Department of Education.

Bilen, Ö., D. Hotaman, Ö. E. Aşkın, and A. H. Büyüklü. 2014. 'LYS Başarılarına göre okul performanslarının eğitsel veri madenciliği teknikleriyle incelenmesi: 2011 İstanbul örneği.' *Eğitim ve Bilim* 39 (172): 78–94.

Boesel, D. 2001. 'Student Attitudes toward High School and Educational Expectations.' Paper presented at annual meeting of the American Educational Research Association, Seattle, 10–14 April.

Christensen, L., R. Johnson, and L. Turner. 2015. *Research Methods Design and Analysis.* Ankara: Anı Yayıncılık.

Christie, C. A., C. M. Nelson, and K. Jolivette, K. 2004. 'School Characteristics Related to the Use of Suspension.' *Education and Treatment of Children* 27 (4): 509–526.

Cohen, L., L. Manion, and K. Morrison. 2002. *Research Methods in Education.* 6th ed. New York: Routledge Taylor and Francis.

Crain-Dorough, M. 2003. 'A Study of Dropout Characteristics and School-Level Effects on Dropout Prevention.' https://search.proquest.com/docview/305320090?accountid=15340

Demirtaş, Z. 2010. 'Liselerde örgüt kültürü ve öğrenci başarısı arasındaki ilişki.' *Mustafa Kemal Üniversitesi Sosyal Bilimler Enstitüsü Dergisi* 7 (13): 208–223.

ERG. 2001. *Çok Programlı Lise Yönergesi.* Ankara: Milli Eğitim Bakanlığı.

ERG. 2005. *Basic Education in Turkey Background Report.* Ankara: Republic of Turkey Ministry of National Education.
ERG. 2011. *Eğitim İzleme Raporu 2011.* Istanbul: Eğitim Reformu Girişimi.
ERG. 2013. *Ortaöğretim Kurumları Yönetmeliği. Resmi Gazete Sayısı: 28758.* Ankara: Milli Eğitim Bakanlığı.
ERG. 2017. 'Egitim Reformu Girişimi: Kamu Bütçesinin Daha Düşük Orani MEB'e Ayriliyor.' https://www.egitimreformugirisimi.org/egitim-reformu-girisimi-kamu-butcesinin-daha-dusuk-orani-mebe-ayriliyor/
ERG. 2018. *2023 Eğitim Vizyonu.* Ankara: Milli Eğitim Bakanlığı.
ERG. 2019. *Eğitimin Yönetişimi ve Finansmanı Eğitim İzleme Raporu.* Istanbul: Eğitim Reformu Girişimi.
García, E., C. Romero, S. Ventura, and C. De Castro. 2011. 'A Collaborative Educational Association Rule Mining Tool.' *The Internet and Higher Education* 14 (2): 77–88.
Grant, J. N. 2016. 'Understanding and Addressing Chronic Absenteeism in Secondary Schools.' https://search.proquest.com/docview/1790627347?accountid=15340
Green, W. L. 2017. 'Educator Technological Pedagogical Content Knowledge and Student Achievement.' https://search.proquest.com/docview/1910117357?accountid=15340
Hair, J., W. Black, B. Babin, and R. Anderson. 2010. *Multivariate Data Analysis.* 7th ed. London: Pearson Education.
Halsell, J. N. 2007. 'Using Cluster Analysis to Evaluate the Academic Performance of Demographic Homogeneous Subsets.' https://search.proquest.com/docview/304764926?accountid=15340
Humphreys, R. G., Jr. 2006. 'The Impact of Institutional Characteristics on Charter School Achievement: The Case of California.' https://search.proquest.com/docview/305369248?accountid=15340
James, G., D. Witten, T. Hastie, and R. Tibshiran. 2013. *An Introduction to Statistical Learning with Applications in R.* London: Springer.
Jermain, B. C. 2018. 'Chronic Absenteeism in Rhode Island Public Schools: A Crisis in the Making.' https://search.proquest.com/docview/2132051179?accountid=15340
Köse, M. 1999. 'Üniversiteye giriş ve liselerimiz.' *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi* 15:51–60.
Kurebayashi, N. 2015. 'Türkiye'de Okullar Arasında Büyük Kalite Farkı Var.' http://www.hurriyet.com.tr/egitim/turkiye-de-okullar-arasinda-buyuk-kalite-farki-var-30189993
Leskovec, J., A. Rajaraman, and J. Ullman. 2014. *Mining of Massive Datasets.* 2nd ed. Cambridge: Cambridge University Press.
MEB. 1999. *Milli Eğitim Bakanlığı Anadolu Liseleri Yönetmeliği.* Ankara: Milli Eğitim Bakanlığı.
Önder, E. 2012. 'İlköğretimde öğrenci başarısında okullar arası eşitsizliklerin analizi.' Doctoral dissertation, Gazi University, Ankara.

Önder, E., and N. Güçlü. 2014. 'İlköğretimde okullar arası başarı farklılıklarını azaltmaya yönelik çözüm önerileri.' *Marmara Üniversitesi Atatürk Eğitim Fakültesi Eğitim Bilimleri Dergisi* 40 (40): 109–132.

Özbay, Ö. 2015. 'Veri madenciliği kavramı ve eğitimde veri madenciliği uygulamaları.' *Uluslararası Eğitim Bilimleri Dergisi* 5:262–272.

Page, B. A. 2004. 'Cluster Analysis: An Alternative Method for Assessing Charter School Student Achievement?' https://search.proquest.com/docview/305168051?accountid=15340

Perry, M. L. 2000. 'Analyzing the Demand for Instructional Personnel in the Virginia Public School System: 1999–2000.' http://search.ebscohost.com/login.aspx?direct=true&db=ddu&AN=9D40EBB189A9682B&lang=tr&site=ehost-live

Shovon, M., H. Islam, and M. Haque. 2012. 'An Approach of Improving Students' Academic Performance by Using K-Means Clustering Algorithm and Decision Tree.' *International Journal of Advanced Computer Science and Applications* 3 (8): 146–149.

Singh, N., and D. Singh. 2012. 'Performance Evaluation of K-Means and Hierarchical Clustering in Terms of Accuracy and Running Time.' *International Journal of Computer Science and Information Technologies* 3 (3): 4119–4121.

Stiefel, L., A. E. Schwartz, and P. Iatarola. 2001. 'Determinants of School Performance in New York Elementary Schools: Results and Implications for Resource Use.' Condition Report Prepared for the Education Finance Research Consortium, New York.

TED. 2008. *80. Yıl Uluslararası Eğitim Forumu: Eğitim Hakkı ve Gelecek Perspektifleri.* Ankara: Türk Eğitim Derneği.

Turnage, I. W. 2011. 'The Relationship of Teacher Efficacy, Teacher Experience, and Teacher Grade Level within the Implementation Process of Behavioral Interventions.' https://search.proquest.com/docview/880398658?accountid=15340

Ungricht, T. R. 1997. 'Learning Strategies of Concurrent Enrolment Students at Utah Valley State College.' https://search.proquest.com/docview/304356644?accountid=15340

Wallace, C. M. 2017. 'Parent/Guardian Perspectives on Chronic Absenteeism and the Factors that Influence Decisions to Send Their Children to School.' https://search.proquest.com/docview/1942891250?accountid=15340

Warren, A. F. 2007. 'The Relationship between Reported Incidents of Student Discipline and Student Achievement Across Four Eastern States.' http://search.ebscohost.com/login.aspx?direct=true&db=ddu&AN=99F77AE26A4F026E&lang=tr&site=ehost-live

Yıldız, M. 2015. 'Türkiye Genelindeki Okullar Arasında Kalite Farkı Büyük.' http://www.hurriyet.com.tr/egitim/istanbul-il-milli-egitim-muduru-yildiz-turkiye-genelindeki-okullar-arasinda-kalite-farki-buyuk-28752799

Sultan Uçar Altun works as a teacher in the field of computer education and instructional technology at a middle school of Antalya province in Turkey. *ealtunsu@gmail.com*

Mualla Aksu works as a professor in the field of educational administration at Akdeniz University in Antalya, Turkey. *muallaaksu@akdeniz.edu.tr*