Original Contribution

ANALYSIS OF ENVIRONMENTAL SENSITIVITY IN THRACE REGION THROUGH TWO STEP CLUSTER

D. Altas[1], A. Kubas[2]*, J. Sezen[3]

[1]Department of Econometrics, Faculty of Economics and Administrative Sciences, Marmara University, Bahcelievler-Istanbul, Turkey
[2]Department of Business Administration, Faculty of Economics and Administrative Sciences, Namik Kemal University, Tekirdag, Turkey
[3]Tekirdag Office of the Minsitry of Envrionment and Urbanisation, Tekirdag, Turkey

ABSTRACT

The study aims to determine the environmental awareness of company managers regarding the solution to various environmental problems caused by industrial enterprises in the region of Thrace. The data were obtained from 110 enterprises through survey technique.

Various documents and information belonging to the enterprises and their related judgment were studied through two- step clustering analysis technique. According to the data obtained, the significance level of TSE, ISO 9000, ISO 14000, and ISO 18000 quality documents owned by the enterprise laws are found to be high. The license for waste water drainage comes second, which is followed by license for emission. In the third cluster, it is seen that managers of the enterprises were sufficiently informed about the basic regulations. This study in general found that although the managers of industrial enterprises acted responsibly, the most important factor here is that the measures to be taken about the environment are the ones that require investment.

Common waste water drainage facilities play an important role in solving the environmental problems experienced in the Thracian region. Enterprises should spare all their time, resources and energy for their area of specialty. A central laboratory and a coordination unit for environmental issues that require expertise should function on behalf of firms and help with solutions to the problems, and firms should finance any eventual expenses.

Key words: Environmental Sensitivity, Cluster Analysis, Two Step Cluster, Industrial Enterprises

INTRODUCTION

After the 1980s, industrial firms have been on rapid increase and they reached the number of 1800. While these firms have positive effects on the socio-economic development of the region, they also cause certain environmental problems. The most important environmental problem is that underground water is drained into the area without being refined adequately after being used in the industrial production.

_____

Correspondence to: *Assoc.Prof.Dr.Ahmet Kubaş, Namik Kemal University, Faculty of Economics and Administrative Sciences, Department of Business Administration, Namik Kemal Mahallesi Kampüs Caddesi No:1, 59030 Tekirdag/Turkey, E-mail: ahmetkubas@yahoo.com Phone number. +90.282.250-2803 Fax number: +90.282.250-9927*

The two most important water resources in the region are the rivers Meric and Ergene flowing through the towns of Saray, Corlu, Muratlı and Malkara of the Tekirdağ province and dividing the Thrace region into two in the middle, where the latter joins the Meric river around the town of Enez with the city of Edirne. All the habitation and industrial companies along the Ergene River drains their waste water off into the river. Called polluted surface water sources, the Ergene River and its branches carry water with both industrial and home waste. For this reason, rivers with excessive overload of pollution become waste water drainage. According to international water quality standards regarding the pollution by industrial

and home waste, the river Ergene has the 5th grade water quality, carrying water that is not suitable for industrial, agricultural and home use. This situation bears risks in terms of environmental and human health.

The rapidly developing industrialization in Thrace brings along environmental problems related to industrialization. The causes, results and solutions of the environmental problems in the area were detected through a questionnaire given to the managers of the industrial businesses. At the interview with managers of randomly selected 110 businesses, issues related to assessments about their approach to environmental problems and whether they hold emission and discharge certificates and EIA (Environmental Impact Assessment) report were examined in terms of the environmental conditions they are in.

To measure environmental sensitivity of the businesses, two ways cluster analysis was applied to the data obtained from the interviews with the managers of the companies. The study aimed to evaluate views and opinions of the managers about the environmental problems in the area and the measurements to be taken within the scope of environmental laws, sanctions and legalities that can be imposed to stop possible waste and rubbish caused by industrial production.

METHOD
1. Cluster Analysis
Cluster analysis is a method that is used to divide variants and units into groups according to their similarities. When we look at the clusters formed as a result of the analysis, we see that while the units in the same clusters look very much alike according to the accepted criteria, the ones in different clusters will look less similar (1). As the homogeneity is increased in the clusters to be formed after the analysis, it is reduced among clusters (2). When the clusters are geometrically placed in the graphic, it is seen that units in the cluster are very close to each other and those in different clusters are differ from each other (3).

The main purpose of the cluster analysis is to group observations by taking distances and similarities into consideration. Apart from this, it has some special goals, such as determining real types, facilitating making up a model that will reduce variants by grouping, forecasting for groups and do testing (4).

As the analysis is more descriptive than predictive and it is concerned with all the relationships in the data set, there is no condition for linearity of the relationships among variants.

Although the normality of the given distribution according to the normal distribution is sought, this hypothesis is not valid in practice. Instead, normality condition of the distance values is sought and this should suffice (4).

Cluster analysis is not used in analysis estimation as it is not dynamic but rather a static analysis method used for describing the current situation. The method has no condition for any scale for data measurement; it is possible to use it in nominal, ordinal, intermittent and rated scales.

One of the most important problems in cluster analysis is to determine the cluster number. Sometimes because of adverse values, a single unit may constitute a cluster. In this situation, there will be at least a few units in one cluster; for instance, issues such as whether single-unit clusters will be accepted or how they will be construed if they are accepted as important in terms of cluster number. Behind all these lies the fact that one must be careful in choosing variants. Although there are different criteria in determining cluster numbers, they don't offer precise solutions in obtaining ideal cluster number; rather, they function as a guide (5, 6).

2. Two Step Clustering
In cluster analysis, there are two different cluster techniques in allocating units or variants according to their similarities. One is hierarchical (7, 8); the other is non-hierarchical (9, 10, 4). In both techniques, the main goal is to bring the differences among clusters and similarities inside the clusters to the highest level. Which technique to use is dependent on the number of clusters that the researcher may accept as being significant. Namely, one of the techniques must be chosen according to the decision making about the number of clusters. In case of predetermination of cluster numbers, non-hierarchical techniques are used.

When you have a really large data set or you need a clustering procedure that can rapidly form clusters on the basis of either categorical or continuous data, none of the previous two procedures fills the bill. Hierarchical clustering requires a matrix of distances between all pairs of cases, and k-means requires shuffling cases in and out of clusters and knowing the number of clusters in advance. The SPSS Two Step Cluster Analysis procedure was designed for such applications. It requires only one pass of data (which is important for very large data files), and it can produce solutions based on mixtures of continuous and categorical variables and for varying numbers of clusters. The clustering algorithm is based on a distance measure that gives the best results if all variables are independent, continuous variables have a normal distribution, and categorical variables have a multinomial distribution (11).

Two way cluster analysis is a technique which can be used in situations when there is no introductory information (12). If the desired number of clusters is unknown, the SPSS Two Step Cluster Component will find the proper number of clusters automatically (13).

Grouping data is realized through Bayesci information and Akaike information criteria (BIC, AIC) in log-likelihood space (Schiopu 2010). These criteria are described below.

$$BIC(J) = -2\sum_{j=1}^{J} \zeta_j + m_J \log(N) \qquad (1)$$

$$AIC(J) = -2\sum_{j=1}^{J} \zeta_j + 2m_J \qquad (2)$$

Here;

$$m_J = J\left\{ 2K^A + \sum_{k=1}^{K^B} (T_K - 1) \right\} \qquad (3)$$

The distance between clusters i and j is defined as

$$d(i, j) = \xi_i + \xi_j - \xi_{<i,j>} \qquad (4)$$

Here,

$$\xi_r = -N_r \left(\sum_{k=1}^{K^A} \frac{1}{2} \log(\sigma_k^2 + \sigma_{rk}^2) + \sum_{k=1}^{K^B} E_{rk} \right) \qquad (5)$$

$d(i, j)$ is the distance between clusters i and j; $<i, j>$ index that represents the cluster formed by combining clusters i and j; K^A is the total number of continuous variables; K^B is total number of categorical variables; Tk is the number of categories for the k-th categorical variable; Nr is the total number of data records in cluster r; $Nrkt$ is the number of records in cluster r whose categorical variable k takes t category; σ_k^2 the estimated variance (dispersion) of the continuous variable k, for the entire dataset; σ_{rk}^2 the estimated variance of the continuous variable k, in cluster r.

Equality (5) E_{rk} is obtained as follows:

$$E_{rk} = -\sum_{t=1}^{T_k} \frac{N_{rkt}}{N_r} \log \frac{N_{rkt}}{N_r}$$

In the literature study it was found that two way cluster analysis could cluster categorical or continuous data together but it was found that it sometimes proved inefficient. However, it was accepted that it could ideally cluster continuous variants (12). Furthermore, in the performed studies it was also believed that in spite of its limitations, gradual cluster analysis could be applied to big data sets that are not homogenous.

ENVIRONMENTAL SENSITIVITY ANALYSIS IN INDUSTRIAL ENTERPRISES

In the study, in order to find out the level of environmental sensitivity in Thrace region, data obtained from the questionnaires applied to 110 company managers were assessed using two way cluster analysis method. Related to the data in the questionnaire, results were assessed through two different analyses. The obtained findings are presented in Tables and graphics below.

1. Two Step Cluster Findings Related to the Problems about the Document

Table 1. *Auto-Clustering*

Number of Clusters	Akaike's Information Criterion (AIC)	AIC Change[a]	Ratio of AIC Changes[b]	Ratio of Distance Measures[c]
1	1365,588			
2	1237,555	-128,033	1,000	1,360
3	1153,978	-83,577	,653	1,203
4	1091,217	-62,761	,490	1,198
5	1045,475	-45,742	,357	1,408

Table 1 shows various cluster numbers obtained to determine suitable cluster number in grouping data according to their similarities. The smallest AIC value (-128,033) according to AIC criteria related to various cluster numbers was obtained for two clusters. Thus, it was found that data should be separated into two clusters.

Table 2. *Cluster Distribution*

		N	% of Combined	% of Total
Cluster	1	74	67,3%	67,3%
	2	36	32,7%	32,7%
	Combined	110	100,0%	100,0%
Total		110		100,0%

Answers given to the questions (K1, K2, K6, K7, K9) of the questionnaire which was prepared to measure the environmental awareness of the industrialist and his/her knowledge about environment were tested by Two Step Cluster Analysis **(Table 2)**. According to the judgments of the 74 business managers who participated in the questionnaire, while 67,3 % of them concentrated on the first cluster, the awareness of the remaining 36 managers (32,7%) concentrated on the on the second cluster. Distribution of the factors in both clusters is as follows.

K1: Quality Certificates the Businesses Hold
Obtaining TSE and ISO Quality certificates is a common goal accepted by all firms. It was realized that private sector has come to a good level, also by knowing that getting these certificates is a must to provide qualified goods and services.

K2: EIA Decision of the Business
EIA (Environmental Impact Assessment) is a process in which the important effects of a project or development on environment are determined. According to EIA regulation published in the Official Gazette of 26939 on 17 July 2008 e 26939, EIA or PRE--Çed decisions are taken according to the lists based on businesses' area of business. PRE- EIA makes the decision that "EIA Effects are not necessary".

K6: Drainage License
In this certificate, waste water parameters of the water discharged into the user's area are

determined. Getting discharge permit is highly important for the businesses in the area.

K7: Type of Fuel Used by the Enterprise
It covers all types of fuels that industrial facilities have to use in any kind of activities (heating, production, etc.). These fuels are coal, natural gas, LPG, Fuel-Oil and others. Natural gas is a gas that is colorless, odorless and lighter than air. (İGDAS) it is the most preferred gas by the industrial businesses as it is a clean fuel. Fuel-oil and LPG are also types of fuels used by businesses.

K9: Emission Allowance License
Air emission and permissions and licenses, that are to be obtained within the context of regulation according to the environment law, are obligatory in order to get environment license needed to establish and operate businesses. This permission is called 'emission allowance certificate'. All the work and operations are done according to the principles and limitations that were outlined in the regulation.
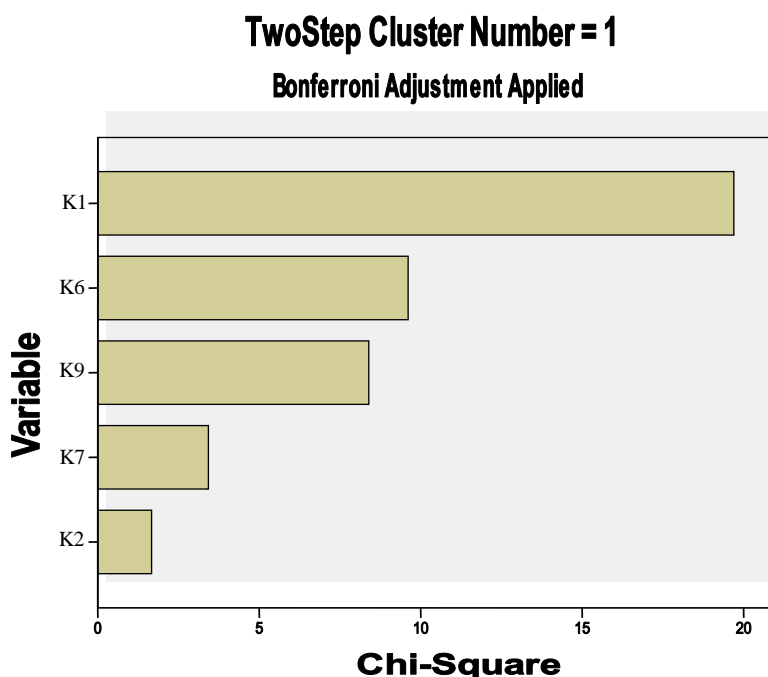
TwoStep Cluster Number = 1
Bonferroni Adjustment Applied

Figure 1. Factors in the First Cluster (Document Questions)

According to **Figure 1**, the significance level of quality certificates that businesses hold was found high in the first cluster of the two step cluster from the questionnaire applied to the industrialists to determine their environmental sensitivity. The discharge permission license is in the second row and emission allowance certificate is in the third row. Type of fuel used in the company is in the fourth row. On the other hand, the EIA decision of the business is in the last row in terms of level of significance according to the analysis done. The reason for this is that the EIA decision is a type of decision that businesses need to make during planning stage. In general, it is inferred that in parallel with the answers given, the obtained quality documents pay an important role in the environmental sensitivity.

According to analysis, quality certificates were found of the highest significance level in the cluster 2 **(Figure 2).** Discharge permission certificate is in the second row; emission allowance certificate is in the third row; type of fuel used in the businesses is in the fourth row; the last row is EIA decision in terms of order of significance.

When both clusters were compared, they display similarities with each other. Type of fuel used in the business and significance level of business's EIA decisions are at the same level in both clusters. First and second cluster distributions

related to "The quality certificates that the businesses hold" variant, which is the most important factor in both clusters, are given in **Figure 3**.
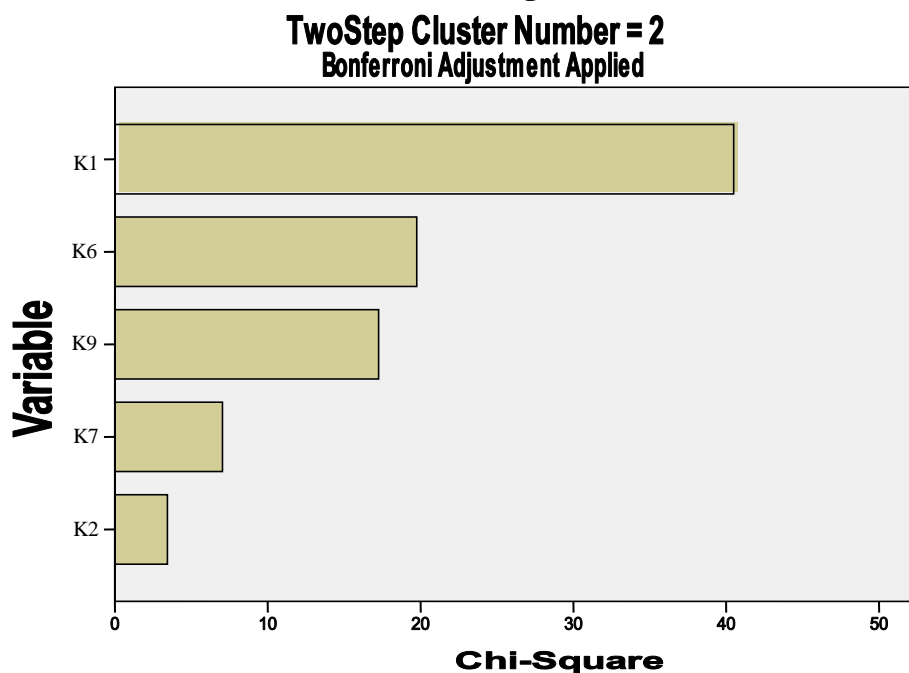
TwoStep Cluster Number = 2
Bonferroni Adjustment Applied

Figure 2. Factors in the Second Cluster
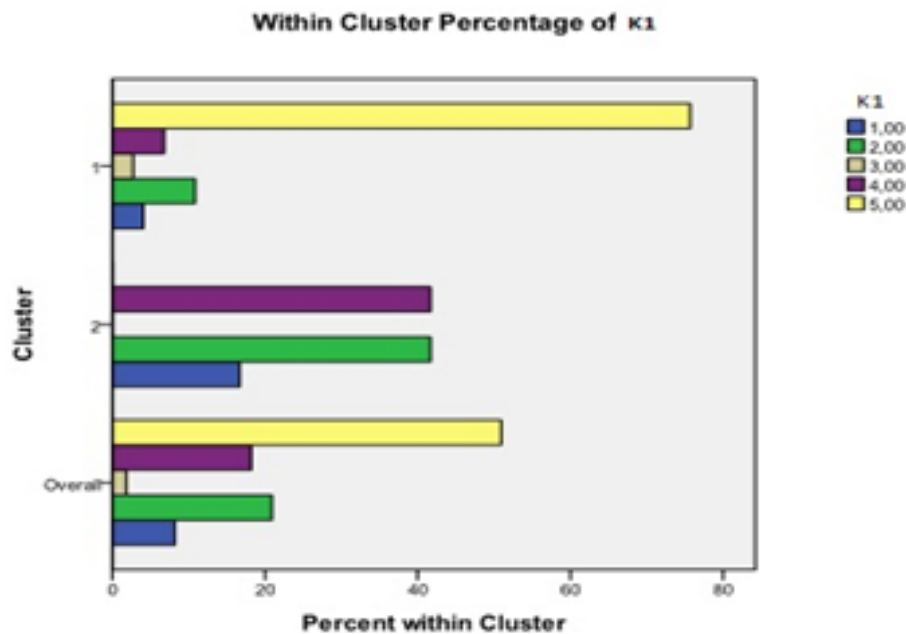
Within Cluster Percentage of K1

Figure 3. Quality Certificates Businesses Hold

Items of the Quality Certificates:
1,00:TSE+ ISO 14000
2,00:TSE+ISO9000+ISO14000

3,00:TSE+ISO18000+ISO14000
4,00:TSE+ISO9000+ISO14000+ISO18000
5,00:Those without a Quality Certificate

Out of the results from the two step cluster analysis, when we look at the rate of the quality certificates in the first cluster (Figure 3), those who don't have a quality certificate are the highest. Those with TSE+ISO9000+ISO14000 quality certificate appear in the first row. According to questionnaire results, those with +ISO9000+ISO14000+ISO18000 were in the third row. Those with TSE+ISO 14000 quality certificates were in the fourth row in the order of questionnaire. On the other hand, for the businessmen, those with TSE+ISO18000+ISO14000 quality certificate were in the last row. In the 2nd group, those with TSE+ISO9000+ISO14000+ISO18000 and

TSE+ISO9000+ISO14000 quality certificates display similarities. It is followed by those with TSE+ISO 14000 quality certificate.

When two clusters were compared with each other, those with no quality certificate are in the first row; businesses possession of quality certificates 2 and 4 is in the second row. For both clusters, those holding TSE+ISO18000+ISO14000 quality certificate are in the last row.

Related to the discharge permission certificate variant which is the second important factor, 1st and 2nd cluster distributions were given in **Figure 4.**
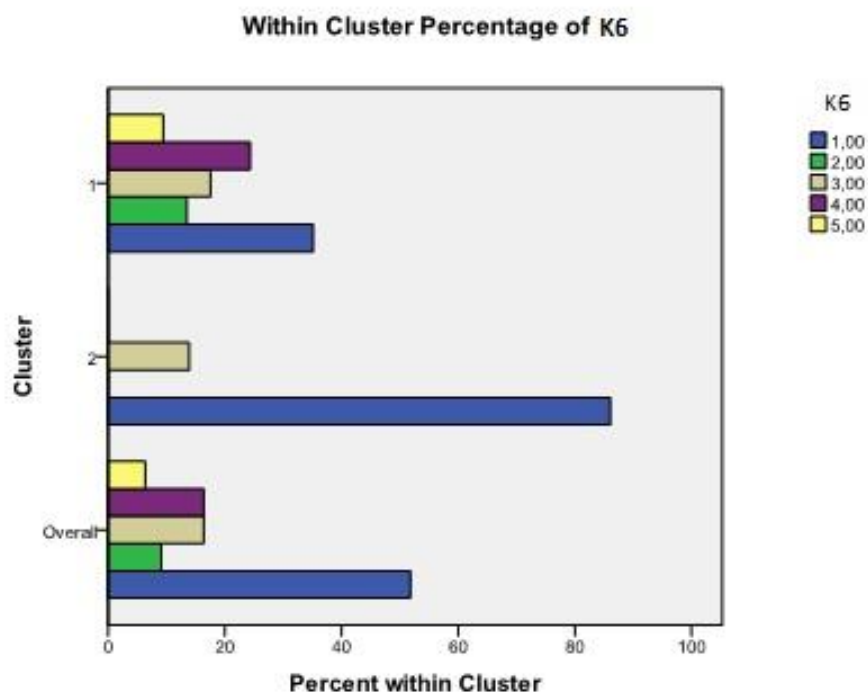

Figure 4. Drainage License

Permission options items for discharge:
1,00: Yes
2,00: No
3,00: I don't need a discharge permission certificate
4,00: In case of application
5,00: I have no idea

Hence, the businesses with discharge permission in the first row are of the highest level of significance. As the water pollution problem is important in the area, it is important to get a discharge permission certificate. The application made by the business for discharge permission is of the second row in order of significance. "No need to obtain discharge permission certificate" option is in the third row in order of significance. Businesses with no discharge permission are in the fourth row in terms of significance but this situation looks insignificant in terms of the area.

The majority of the businesses with no discharge permission, certificate is in the last row in terms of significance.

Businesses with discharge permission certificate are in the second cluster are of the highest level of significance. Businesses that do not possess a discharge permission certificate are in the second cluster, while other options were not available in order of significance. When we look at the whole of two clusters, while the first row belongs to the businesses with discharge permission certificate,

the second row is occupied by the businesses that don't have to discharge and those who applied for discharge permission certificate have the same level of significance. While the businesses with no discharge permission certificate are in the third row at the level of significance, businesses that have no idea about the discharge permission are in the last row.

2. Two Step Cluster Findings Related to the Questions about Knowledge

Table 3. *Cluster Distribution*

		N	% of Combined	% of Total
Cluster	1	53	48,2%	48,2%
	2	57	51,8%	51,8%
	Combined	110	100,0%	100,0%
Total		110		100,0%

Answers given to the questions, if they have knowledge about K9:Emission permits; K10:Environmental control; K11:Recyclying; K16:Hazardeous wastes; K19:Regulations; K20:Environmental pollution; K21:Noise pollution; K22:Waste water, have been tested

through two-step cluster analysis method **(Table 3)**. According to the answers given by 53 business managers who filled the questionnaire, while 48, 2 % concentrated on the first cluster, the judgments of the rest 57 (51, 8 %) concentrated on the second cluster.
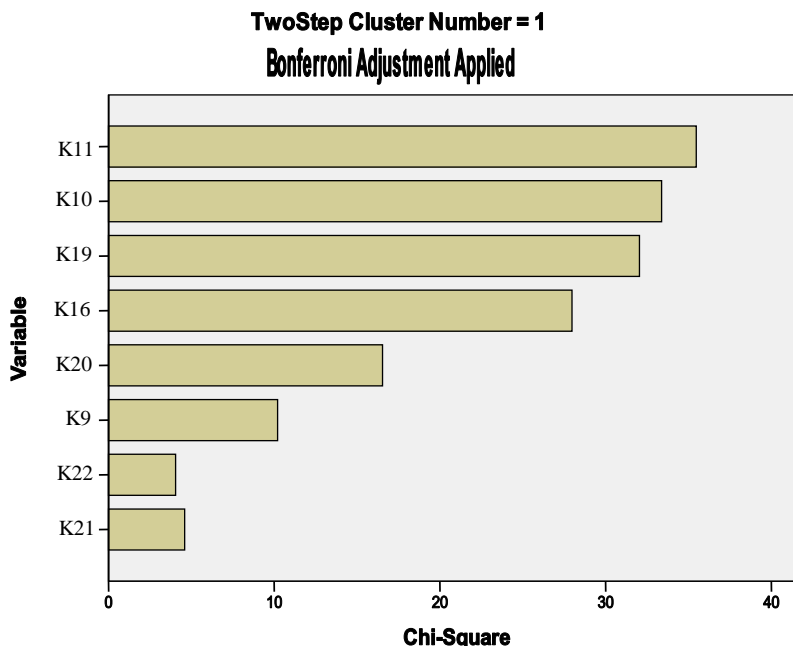

Figure 5. Factors in the First Cluster (Knowledge Questions)

According to **Figure 5**, as a result of the analysis of the questionnaire, businessmen's knowledge about recycling is in the first row in terms of

environmental sensitivity. It is clear that businesses of the Thrace region aim at evaluating all kinds of waste and pollutants resulting from

their production. As in the most of developed countries, recycling is highly important for businessmen and it also provides raw material for other industrial activities. In the second row of significance comes businesses knowledge about environmental control. According to analysis result, in the third row, businesses are responsible for following the Environmental law 2872 and all the related regulations. Information about dangerous waste caused by businesses is in the fourth row. Dangerous wastes that can be caused by businesses are subject to the limitations of the regulation and they require highly binding applications and procedures for the businesses. Participation in panel discussion,

seminars and meetings about environmental pollution and problems is in the fifth row and it has lower rate of significance compared to other factors. Emission allowance certificate is in the sixth row. The reason for the presence of emission allowance certificate in this group among the questions in the questionnaire is that air pollution affects all the towns and all the factories are part of this problem. According to analysis results, precautions against noise caused by the business with little difference are in the seventh row and finally question about the solution to the water pollution in the area is in the last row.
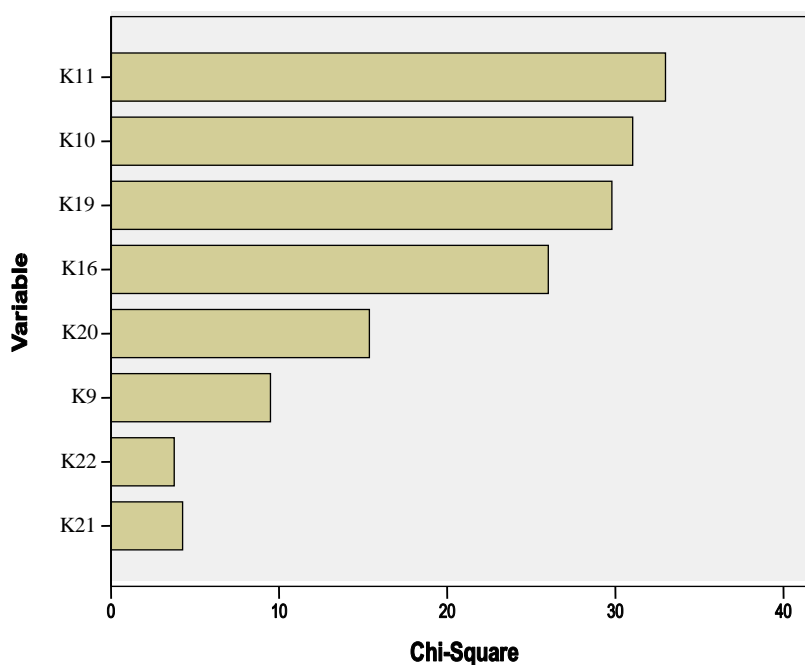

Figure 6. Factors in the Second Cluster (Knowledge Questions)

Hereby, as mentioned above, recycling is an area that businesses are highly sensitive about. According to the analysis, environmental supervision remaining behind sensitivity level and being followed by the Environmental law 2872 and related regulations as in cluster 1 is another issue that businesses are sensitive about. According to the analysis done, other results in the second cluster follows as environmental pollution, rate of participation to activities such

as panel discussion, seminars or meetings about environmental issues, availability of emission allowance certificate, precautions that can be taken against noise caused by the business and solution to the water pollution.

First and second cluster distributions related to k11 variant that is the most important factor are given in **Figure 7**.
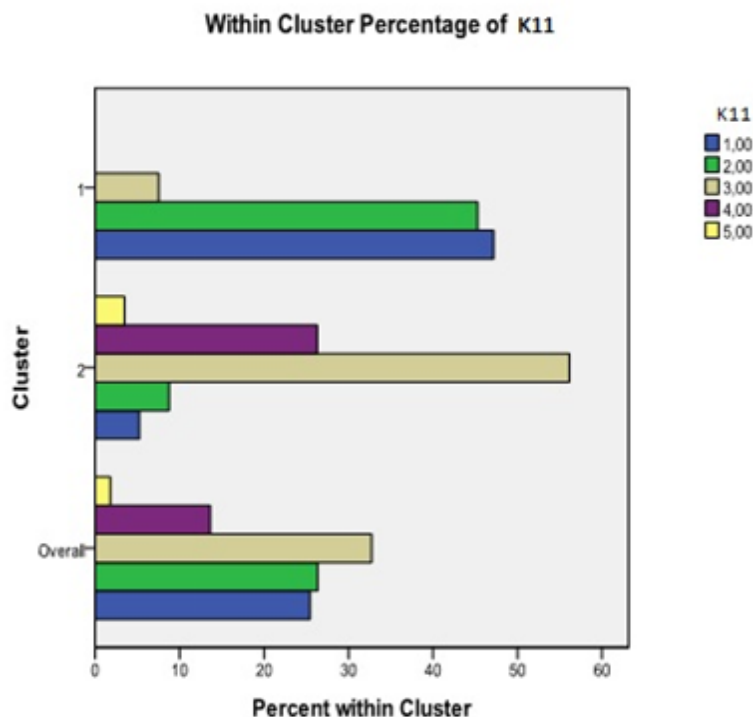
326

Within Cluster Percentage of K11

Figure 7. k11

1,00: Definitely adequate
2,00: Rather adequate
3,00: Adequate
4,00: Not adequate
5,00: Not available

According to **Figure 7** that shows the knowledge of businessmen about recycling, the option "adequate" is in the first row showing that businessmen are sensitive about recycling. The "rather adequate" option in the second row shows sensitivity with little difference compared to the first option. Eliminating "adequate" option in the last row, options of "not adequate" and "not available" about recycling are redundant in the first row. Eliminating "adequate" option in the last row, options of "not adequate" and "not available" who suggest the judgment of insignificance were not available in cluster 1. In cluster 2, the option that businessmen's knowledge about recycling was found to be sufficient. In the last row it was the option that there was no information about recycling.

When two clusters are compared with each other regarding the order of significance, the first row belongs to the option that one is informed enough about recycling. In the second row comes "rather adequate" option, which is followed by "definitely adequate" option in the third row. In the fourth row, there is "not adequate" option. In the last row, there is "not adequate" option and the last row includes "businessmen don't have enough information about recycling".

According to these results, it is convincing enough that quite a considerable number of businessmen who do business in various areas in Thrace region and produce a lot of waste are not sensitive enough about recycling.

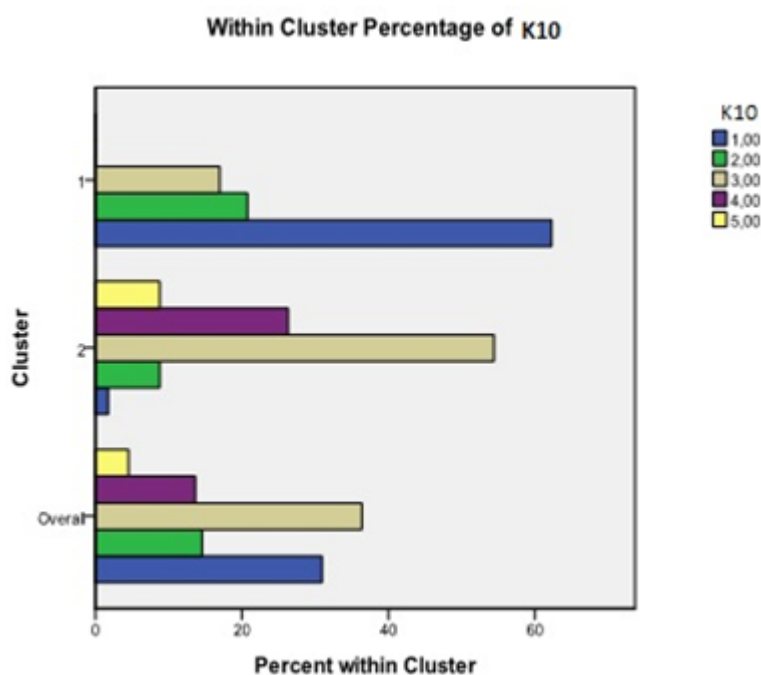Results for k10, which is the second important factor, are as follows:

Within Cluster Percentage of K10

Figure 8. k10

1,00: Definitely available
2,00: Rather available
3,00 Available
4,00: Not much
5,00: Not at all

According to **Figure 8** which shows the business environmental supervision, the cluster 1 shows that businessmen are definitely informed about environmental supervision and that they are also highly sensitive. The second row shows that businessmen are highly informed about environmental supervision. The third row shows the option that businessman is informed about supervision. Fourth and fifth options which are related to the lack of information about supervision were not available in cluster 1. In the results from cluster 2 analyses, which are different from cluster 1, the option that the highest level of sensitivity was related to businessmen's knowledge about environmental supervision. The second row shows that businessmen have little information about their supervision. About the supervision, the third row shows the option "very much informed" and "no information at all". The last row shows that businessmen are definitely informed about supervision. Businessmen's sensitivity about environmental supervision in the second cluster is considerably lower than that of in the first

cluster. This shows that businesses in some areas of the region are more sensitive to supervisions to avoid legal sanctions and some were superficially interested in supervision.

In the combination of both clusters, it is understood that the highest rate goes to "available" option 3, showing that sensitivity of the businessmen in the area to supervision is on the average. "Definitely available" option in the second cluster with precise emphasis on sensitivity is the average rates that businessman in the first cluster shows proportionally. "Rather available with little difference" option is in the third row and "available not much" option is in the fourth row. The option that "there is no knowledge about supervision" is the last row of the analysis that shows the combination of two clusters.

RESULTS AND SUGGESTIONS
As a result of two step cluster analysis applied to industrialists' questionnaire that was targeted at determining environmental sensitivity, primarily the data related to the questions about certificates obtained were evaluated. Knowing that the number of clusters which are appropriate according to AIC criteria is two, questions in this cluster were examined according to the order of importance. The order of significance of the

quality certificates was found high. Waste water discharge certificate is in the second row, emission allowance certificate is the third row. While the type of fuel used in the business is in the fourth row, it was realized that industrial businesses exert extra effort about quality and standards of waste water discharge and emission permission.

Questions prepared to measure businessmen's environmental sensitivity and their knowledge about environment were studied using two step cluster analysis and cluster number was decided to be two. The orders of significance of the judgments were found to be the same. When cluster results were studied, businessmen's knowledge about recycling of fossil waste in terms of environmental sensitivity is in the first row and the level of environmental supervision sensitivity is in the second row. According to the analysis done, the other results in the clusters are as follows: following all the environmental regulations issued, information about the dangerous waste caused by the business, participation rate to panel discussions, seminars, meetings, etc. about environmental pollution or other environmental problems, possible precautions that can be taken against noise caused by the business and solution to water pollution in the region.

It is seen that business managers of the industrial businesses and the people of the Thrace area are informed enough about environmental problems and their solutions. Results obtained from interviews with industrial businesses also support this fact. It was also found that parties have coordination and organization problems in solving environmental problems in the area. As the new generation of industrialists is more sensitive about environment, current surface water pollution problems in the region can be solved by establishing common refinement facilities.

One should benefit from the experience and knowledge of the managers of the businesses regarding the environmental problems experienced in Thrace region. Businesses can gather together and work for the realization of

activities such as common refinement facilities, accredited chemical and biological analysis laboratory or recycling store. Solving environmental problems on the spot is very essential due to the productivity of the businesses.

REFERENCES

1. Chatfield,C.; Collins, A.: Introduction to Multivariate Analysis, London and Newyork, 1980.
2. Zingmund, W. G.: Exploring Marketing Research, Fifth Edition, The Dryden Press, Orlando, 1994.
3. Mcdaniel, J.C., Gates, R.: Contemporary Marketing Research, Third Edition, West Pub. Co., San Francisco, 1996.
4. Tatlıdil, H.: Uygulamalı Çok Değişkenli İstatiksel Analiz, Ankara, 2002.
5. Gürbüz, H., Yıldız, V., Yıldız,Z.: Ülkemizin Turizm Stratejilerinin Planlanmasında Kümeleme Çözümlemesinin Kullanılması, IV. National Econometrics and Statistics Symposium (Antalya 14-16 May 1999), Marmara University, Faculty of Economics and Administrative Sciences Pub., İstanbul, 2000.
6. Özdamar, K.: Statistical Data Analysis with Package Programs-II, Kaan Pub., Eskişehir 1999.
7. Güler, F., Ulutürk, S.: Kümeleme Analizi ve Bir Uygulama, IV. National Econometrics and Statistics
1. Symposium (Antalya 14-16 May 1999), Marmara University, Faculty of Economics and Administrative Sciences Pub., İstanbul, 2000.
8. Johnson, R. A., Wichern, D.W.: Applied Multivariate Statistical Analysis, Prentice-Hall Inc., Fifth Edition, New Jersey (2002)
9. Manly, B.F.J.: Multivariate Statistical Methods A Primer, Chapman and Hall, London, 1986.
10. Anderberg, M.: Cluster Analysis For Applications", New York, 1973.
11. http://www.norusis.com/pdf/SPC_v13.pdf
12. Kayri,M.: Arastırmalarda İki Asamalı Kumeleme (Two-Step Clustering) Analizi ve Bir Uygulaması, Eurasian Journal of Educational Research, 7(28), 89-99, 2007.
13. The SPSS Two-Step Cluster Component,White paper–Technical report. www.spss.ch/upload/1122644952_The%20SPSS%20TwoStep%20Cluster%20Component.pdf