Classification of Turkish Commercial Banks Under Fuzzy c-Means Clustering

İsmail Hakkı GÖKGÖZ*
Fatih ALTINEL**
F. Pınar Yetkin GÖKGÖZ***
İlker KOÇ****

Abstract

As the major actors of credit system, banks have a great importance not just for financial system but also for the whole of economy. Thus, financial soundness of banks, affected by many financial risks, should be monitored closely. This study focuses on classification of the deposit and participation banks of Turkey regarding their soundness. Financial Stability Indicators (FSIs) are used to attain this goal. Research method is mainly based on fuzzy c-means clustering method which relies on fuzzy logic. The results show that the participation banks are grouped together in the same cluster. Also, Denizbank A.Ş., Finansbank A.Ş., Yapı ve Kredi Bankası A.Ş. and Türk Ekonomi Bankası A.Ş., having similar characteristics regarding ownership and scope of financial services, are found to be grouped together in all periods under consideration. Moreover, it has been seen that size is not the most decisive factor for classification purposes.

Key Words: *Financial Risk, Financial Soundness Indicators, Turkish Commercial Banks, Data Clustering, Fuzzy c-Means Clustering.*

JEL Classification: *C38, C61, G24, G32*

Özet - Türk Bankacılık Sistemindeki Bankaların c-Ortalamalı Bulanık Kümeleme Analizi ile Sınıflandırılması

Kredi sisteminin ana aktörlerinden olan bankalar sadece finans sistemi için değil tüm ekonomi için büyük önem taşır. Bu sebeple, birçok riskle karşı karşıya olan bankaların sağlamlıklarının yakından izlenmesi gerekir. Bu çalışmada, Türk mevduat ve katılım bankalarının finansal sağlamlıklarına göre sınıflandırılması amaçlanmıştır. Bu amaç için Finansal Sağlamlık Göstergeleri (FSIs) kullanılmıştır. Çalışma ana yöntem olarak c-ortalamalı bulanık kümeleme analizine dayanmaktadır. Çalışma sonucunda katılım bankalarının birlikte gruplandığı görülmüştür. Bunun yanında, sahiplik ve faaliyet gösterilen alan açısından benzer özellikler taşıyan Denizbank A.Ş., Finansbank A.Ş., Yapı ve Kredi Bankası A.Ş. ile Türk Ekonomi Bankası A.Ş.'nin araştırmaya konu olan tüm dönemlerde aynı grup altında gruplanmıştır. Ayrıca, finansal büyüklüğün gruplandırmada en belirleyici gösterge olmadığı sonucuna ulaşılmıştır.

Anahtar Kelimeler: *Finansal Risk, Finansal Sağlamlık Göstergeleri, Türk Ticari Bankaları, Veri Kümeleme, c-Ortalamalı Bulanık Kümeleme.*

JEL Sınıflandırması: *C38, C61, G24, G32*

* Banking Specialist, Banking Regulation and Supervision Agency
** Banking Specialist, Banking Regulation and Supervision Agency
*** Treasury Specialist, Undersecretariat of Treasury
**** Banking Specialist, Banking Regulation and Supervision Agency
The views expressed in this paper are solely of the authors and do not necessarily reflect the views of the Banking Regulation and Supervision Agency and Undersecretariat of Treasury.

1. Introduction

The overall health of an economy is directly related with the soundness of financial markets. In that sense, determination of the fragility of banks which are the main actors of financial system is a crucial concern that regulators should closely deal with. Turkish banking sector experienced a great depression in the year 2001 and in the following periods, yet has not faced with any distressed time horizon except from İmar Bank fraud which occurred in 2003. Besides the strengths of Turkish economy, there was another reason that protected Turkish financial system from hazardous effects of global crisis that took place in 2008. The reason for that was the very low investment level of Turkish banks in mortgage-backed securities especially structured by Lehman Brothers. Thus, failure of Lehman did not directly affect Turkish banks compared to the banks of European economies struggling in financial crisis. Therefore, being far from distressed conditions for a long time horizon and living in a financially troubled world makes the evaluation of Turkish banks more critical and crucial.

A major objective of this paper is to classify commercial Turkish banks (Deposit banks and participation banks[1]) according to their credit qualities. For that purpose Financial Soundness Indicators (FSIs) of banks have been used (International Monetary Fund, 2006). These ratios have been calculated by using the data obtained from databases of The Bank Association of Turkey and from the web sites of participation banks. Data set consists of three periods: March 2012, June 2012 and September 2012, details of which are explained in Section 4. The organization of the paper is as follows: The short introduction in Section 1 is followed by the literature review in Section 2. In Section 3, firstly, the basics of data clustering, hard clustering method which relies on classic set theory and also fuzzy clustering method which has been derived from fuzzy logic and fuzzy sets, are described. Secondly, fuzzy c-means clustering method and steps of fuzzy c-means clustering algorithm have been explained. An implementation of fuzzy c-means algorithm to data regarding FSIs of banks is described in Chapter 4. Finally, conclusion includes conclusive arguments and a short summary of this work.

[1] Islamic banks

2. Literature Review

The evaluation of the creditworthiness of an institution is a former issue, which academicians and practitioners have studied. And data clustering is one of the numerous methods, which are used to evaluate the creditworthiness of borrowers. Fundamentally, the process for appraisal of the credit quality is the process of default estimation for the firm under consideration. The methods used to measure default risk of institutions generally rely on qualitative tools, quantitative tools or both. Readers can refer to Servigny and Renault (2004), OeNB and FMA (November 2004) and Gökgöz (2012) for detailed information about credit assessment models that are used to evaluate credit quality of issuers or borrowers.

In literature, the number of studies on clustering algorithms applied on the Turkish banking sector data or on the whole financial market data in a broader sense is limited. The study, applied to financial ratios of companies quoted to Istanbul Stock Exchange (ISE), worked by Tufan and Hamarat (2003) is one of the studies which relies on fuzzy clustering method. Regarding the banking data, Doğan (2008) classifies Turkish banks by means of hard clustering techniques. However, his study does not include any information about fuzzy clustering techniques.

In international literature, fuzzy clustering methods are generally used to segregate credit quality of the commercial loans. One of the main studies in this area performed by Chen and Chiou (1999) to segregate credit quality of Taiwan commercial loan's customers. This study uses the fuzzy integrals in order to split commercial loans regarding their credit qualities. The article written by Alam, Booth, Lee, & Thordarson (2000) is one of the sparse studies trying to predict banks' failure by means of fuzzy clustering and neural networks .

Regarding its methodology, data and scope, this study is one of the novel studies which concantrate on the Turkish commercial banks data. Compared to other clustering methods such as hard clustering tecniques, fuzzy c-means clustering has fundemantal advantages on banking and financial data, because of its great similarity to human decison making ability. These properties make this study very intriguing.

3. Data Clustering

Cluster analysis is one of the basic techniques used for classification of data. In the literature, this technique is used for various areas such as geology, medicine and engineering systems etc. (Yang, 1993). Generally, clustering algorithms are not constructed on common statistical methods; they do not rely on the distribution of the underlying data. Thus, clustering methods are useful when sufficient prior knowledge does not exist.

In general, clustering analysis refers to a wide range of methods which try to divide a data set A in to c subsets (Bezdek, Ehrlich, & Full, 1984).

Let $A = \{a_k | k = 1,2, \ldots, N\}$ be a sample of observations in n-dimensional Euclidian space $\mathbb{R}^n$, a_k is the k-th feature vector; a_{kj} is the j-th feature of a_k. That is, each observation has n measured variables which are represented by column vector $a_k = \left[a_{1k}, a_{2k,\ldots,} a_{nk}\right]^T$, $a_k \in \mathbb{R}^n$. Thus the data set $A = \{a_k | k = 1,2, \ldots, N\}$ with N observations can be represented by an $n \times N$ matrix (Babuska, 2009).

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nN} \end{bmatrix} \qquad (1)$$

The columns of the data (pattern) matrix A are called patterns or objects and the rows of it are called the features or attributes in general pattern-recognition terminology. In a different context, the columns and rows of matrix A may be transposed. For example, the columns of A may represent the patients whereas the rows may represent the symptoms. In the context of this study, columns will represent banks and rows will represent the financial ratios of banks.

There are various definitions of a cluster, depending on the aim of clustering. Clustering is the process of grouping the most similar data into different classes or clusters. Therefore, meaning of the term "similarity" has to be defined. According to Bezdek (1981) similarity should be taken as mathematical similarity, measured in a well-defined sense. Generally, the dissimilarity between two patterns is defined on the feature space by use of a distance metric. As well as distance can be measured among the data vectors themselves, it can also be measured between data vectors and prototypes of the cluster. These prototypes are not known before the application of the clustering algorithm, rather they are sought by the clustering algorithm during process of clustering. These prototypes can be vector of same dimensions like real pattern vectors, but they can also be higher-level linear or non-linear geometrical objects (Babuska, 2009).

Definition 2.1: A metric space is an ordered pair $(\mathcal{M}, d)$ where $\mathcal{M}$ is the set and

$d: \mathcal{M} \times \mathcal{M} \rightarrow \mathbb{R}$ is the function which satisfies the following conditions for any $x, y, z \in \mathcal{M}$:

1. $d(x, y) \geq 0$
2. $d(x, y) = 0 \; iff \; x = y$
3. $d(x, y) = d(y, x)$
4. $d(x, z) \leq d(x, y) + d(y, z)$

There are many types of distance functions (metric) in case of continuous features (variables). Because of the fact that the distance functions have different geometries, each of these functions implies a different view of data (Pedrycz, 2005). For example, if Euclidean metric is used, then data shape will be circular. Table 1 shows some of the famous distance metrics.

Table 1: Some distance metrics between two patterns x and y (Pedrycz, 2005)

Distance Metric	Formula
Euclidian Metric	$d(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$
Hamming Metric	$d(x, y) = \sum_{i=1}^{n} \lvert x_i - y_i \rvert$
Tchebyschev Metric	$d(x, y) = \max_{i=1,2,\ldots,n} \lvert x_i - y_i \rvert$
Minkowski Metric	$d(x, y) = \sqrt[p]{\sum_{i=1}^{n} (x_i - y_i)^p}$

The geometry of distance metrics can be illustrated easily just by taking two patterns $x = [x_1 x_2]^T$ and $y = [00]^T$ with only two features. Figure 1 shows the Euclidean and Hamming distance geometries in three dimensional space.

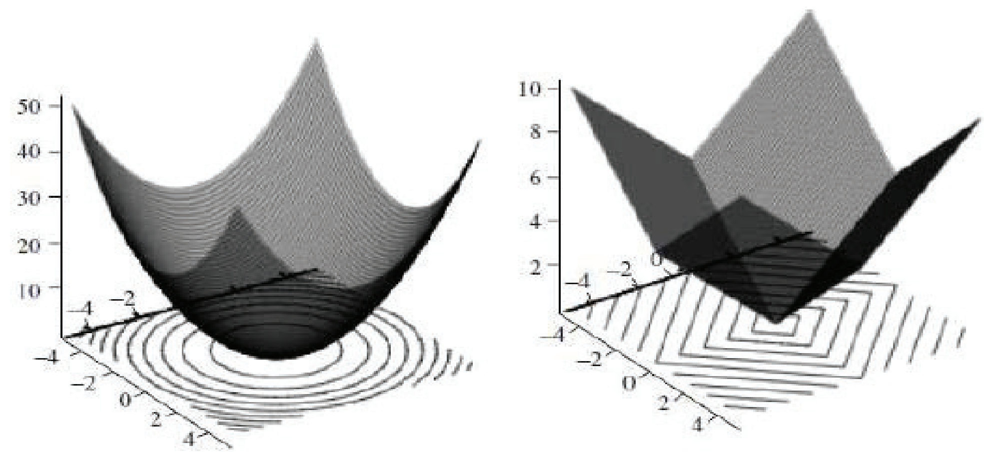


Figure 1: Three dimensional Euclidean and Hamming distance metrics (Pedrycz, 2005)

Although there are a lot of clustering algorithms in literature, in the context of this study, only two of the main clustering methodologies will be mentioned, namely *hard* (crisp) or *fuzzy* clustering.

Hard clustering methods restrict the model that each point of datum is assigned to a single cluster (Yang, 1993). In other words, this clustering method is based on classical set theory; therefore an object can only be member of single cluster or set.

Fuzzy clustering methods are based on *Fuzzy sets* and *Fuzzy logic* which are developed by Zadeh (1965) in order to solve problems related to the pattern classification and cluster analysis.

3.1. Hard Clustering

Let A be a subset of an n-dimensional Euclidean space $\mathbb{R}^n$ with Euclidean norm $\|.\|$ and let c be a positive integer having a value larger than 1. Under hard-partition the aim of clustering is to partition A into c classes. A hard-partition of the set A into disjoint subsets. $A_1, A_2, \dots, A_c$ using classical set theory should satisfy following conditions (Bezdek, 1981):

$$A_1 = \bigcup_{1 \leq i \leq c} A_i \qquad (2)$$

$$A_i \bigcap A_j = \emptyset, \qquad 1 \leq i \neq j \leq c, \qquad (3)$$

$$\emptyset \subset A_i \subset A, \qquad 1 \leq i \leq c. \qquad (4)$$

A partition of a set can also be described by recasting conditions $A_1 = \bigcup_{1 \leq i \leq c} A_i$ in matrix form (Bezdek, Ehrlich, & Full, 1984). By using membership (characteristic) functions u_{ik}, the partition matrix $U = [u_{ik}]_{c \times N}$, U is the matrix representation of hard-partition of A if followings are hold:

$$u_k(a_k) = u_{ik} = \begin{cases} 1 \; if \; a_k \in A_k \\ 0 \; if \; a_k \notin A_k \end{cases} \qquad (5)$$

$$\sum_{i=1}^{N} u_{ik} = 1, \qquad 1 \leq k \leq N, \qquad (6)$$

$$0 \leq \sum_{i=1}^{N} u_{ik} \leq N, \qquad 1 \leq i \leq c, \qquad (7)$$

The space of all possible hard partitions of A under hard-partition methodology has the following representation (Bezdek, 1981):

$$\mathcal{M}_{hc} = \left\{ U \in \mathbb{R}^{c \times N} \middle| u_{ik} \in 0,1, \forall i, k; \sum_{i=1}^{N} u_{ik} = 1, \forall j; 0 \leq \sum_{i=1}^{N} u_{ik} \leq N, \forall i \right\} \qquad (8)$$

The problem with hard-partition methods is the constraint that one element or data can be a member of only a single class or cluster. However, when we try to cluster the data set $A = \{a_k | k = 1,2, \dots, 10\}$, which has a form like the one given in Figure 2, in two dimensional Euclidian space $\mathbb{R}^2$ into two clusters, it is difficult to assign the data points a_9, which is in the middle of the two clusters, and the "outlier" data point a_1 to first or second cluster. Under hard-partition, the data point a_9 is either an element of the first or the second cluster. However, since it is located in between both sets, it should have some probability to be member of each set. Since it is closer to set 1 (upper left corner of the Figure 2) it should have greater probability of belonging to set 1 or in other words, membership function value of a_9 for set 1 has a higher value than that for set 2.

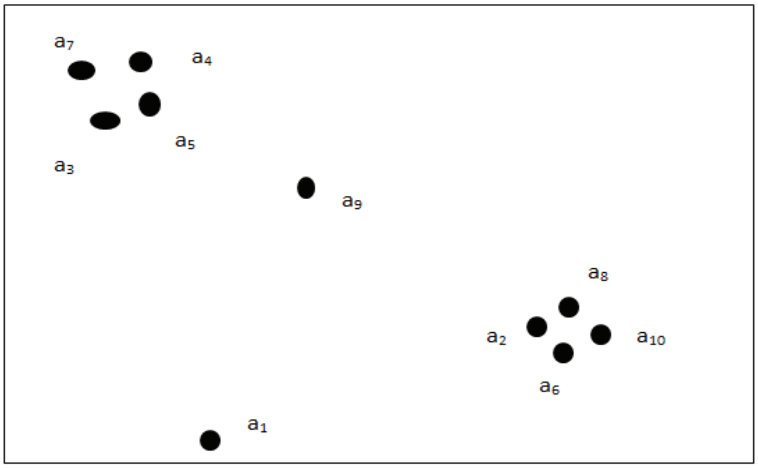


Figure 2: A data set in two dimensional Euclid space.

A partition of data set $U \in \mathcal{M}_{hc}$ in Figure 2 under hard-partition methodology has the following matrix representation:

$$U = \begin{Bmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{Bmatrix}$$

The partition structure given by the matrix U classifies the data point a_1 in first cluster and a_9 in second class, but a_9 also has same positive Euclidean distance to first cluster. Therefore, this partition cannot be considered as optimal partition. To overcome these weaknesses of hard-clustering methods, fuzzy-clustering methods have been proposed in the work of Bellman et. al. (1996) and Ruspini (1969).

3.2. Fuzzy Clustering

Fuzzy clustering is based on fuzzy set theory and it is a generalization of the hard-clustering method just by allowing membership function u_{kj} to take continuous values in the interval [0,1] (Bellman, Kalaba, & *Zadeh* 1996). According to Ruspini (1969) the space of all matrix partitions of data set A, has the following form under fuzzy set theoretical approach:

$$\mathcal{M}_{hc} = \left\{ U \in \mathbb{R}^{c \times N} \middle| u_{ik} \in [0,1], \forall i, k; \sum_{i=1}^{c} u_{ik} = 1, \forall k; 0 \leq \sum_{i=1}^{N} u_{ik} \leq N, \forall i \right\} \quad (9)$$

Conditions for fuzzy-partition matrix may be restated more explicitly as follows:

$$u_{ik} \in [0,1], 1 \leq j \leq c, \qquad 1 \leq k \leq N, \qquad (10)$$

$$\sum_{i=1}^{c} u_{ik} = 1, \qquad 1 \leq k \leq N, \qquad (11)$$

$$0 \leq \sum_{k=1}^{N} u_{ik} \leq N, \qquad 1 \leq i \leq c, \qquad (12)$$

The values of the ith row of the fuzzy partition matrix U are ith membership function of the fuzzy set A_i of A (Babuska, 2009). As it has been stated in equation (10), the sum of every columns of U is equal to 1. One of the possible fuzzy-clustering of data set in Figure 2 may have the following matrix representation:

$$U = \begin{Bmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0.7 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0.3 & 1 \end{Bmatrix}$$

There are different kinds of fuzzy-clustering methods in literature. Probabilistic partition introduced by Krishnapuram (1993) is one example of fuzzy clustering and its main difference is that it relaxes the assumption of "sum of each column of partition matrix U to be equal to 1". Instead of $\sum_{i=1}^{N} u_{ik} = 1,\ 1 \leq k \leq N,$ this method assumes that $u_{ik} \geq 0$. That is, under this method $\sum_{i=1}^{c} u_{ik}$ can be less than 1. For example, under probabilistic fuzzy clustering, data point a_9 would be member of each class with relative probabilities 0.3 and 0.1 and sum of membership function values of both classes equals to 0.4 instead of 1.

3.2.1. Fuzzy c-Means Clustering

Besides being one of the most widely used fuzzy clustering algorithm, fuzzy c-means clustering (FCM) is one of the objective function based clustering methods. The data partition is done by minimizing the fuzzy c-means functional introduced by Dunn (1973).

$$\mathcal{H}(A;\mathbb{U},\mathbb{V}) = \sum_{i=1}^{c}\sum_{k=1}^{N}(u_{ik})^m \|a_k - v_i\|_A^2 \tag{13}$$

where $\mathbb{U} = [u_{ik}]$ is matrix representation of data set A, $\mathbb{V} = [v_1, \dots, v_c]$, $v_i \in \mathbb{R}^n$ is a prototype (centers) vector of cluster that has been determined during process of algorithm run, $d_{ik}^2 = \|a_k - v_i\|_A^2 = (a_k - v_i)^T A(a_k - v_i)$ is the squared inner-product distance norm, and $m \in [1, \infty)$ is a fuzziness parameter of the resulting cluster. The objective function (13) is also called cost function. The minimization of the cost function can be seen as the minimization of the variation between data points vector a_k and prototype vector v_i (Babuska, 2009). As it is observed FCM algorithm uses Euclidean metric as the dissimilarity measure. The optimization of functional $\mathcal{H}(A;\mathbb{U},\mathbb{V})$ is a non-linear optimization and can be solved by various methods. If Picard iteration is used then the membership matrix and centers vector can be calculated as follows:

$$u_{ik} = \frac{1}{\sum_{j=1}^{c}\left(\frac{\|a_k - v_i\|_A^2}{\|a_k - v_j\|_A^2}\right)}, 1 \le i \le c, 1 \le k \le N, \tag{14}$$

$$v_i = \frac{\sum_{k=1}^{N}(u_{ik})^m a_k}{\sum_{k=1}^{N}(u_{ik})^m}, 1 \le i \le c, \tag{15}$$

3.2.2. Fuzzy c-Means Clustering Algorithm

In order to cluster data set A into c number of clusters $(1 \le c \le N)$ (c is initially chosen), with fuzziness exponent $(1 \le m < \infty)$ and the termination tolerance $\varepsilon > 0$, the partition matrix $\mathbb{U}_0$ should be initialized. General algorithm steps are given below:

Step 1: Initialize partition matrix $\mathbb{U}_0$,

Step 2: Calculate centers (mean) vectors at each step:

$$v_i = \frac{\sum_{k=1}^{N}(u_{ik})^m a_k}{\sum_{k=1}^{N}(u_{ik})^m}, 1 \le i \le c$$

Step 3: Compute distances and update the partition matrix $\mathbb{U}(j)$, $\mathbb{U}(j+1)$

$$u_{ik} = \frac{1}{\sum_{j=1}^{c}\left(\frac{\|a_k - v_i\|_A^2}{\|a_k - v_j\|_A^2}\right)}, 1 \leq i \leq c, 1 \leq k \leq N,$$

Step 4: Stop if $\|\mathbb{U}(j+1) - \mathbb{U}(j)\| \leq \varepsilon$, otherwise go to step 2.

4. Data Description and Application

4.1. Data Description

In this study the FSIs of Turkish commercial banks for the periods March, June and September 2012 are used. FSIs, which are widely used for estimation of financial soundness of institutions, have two categories; core indicators and encouraged indicators. Core FSIs consist of twelve ratios which focus on capital adequacy, profitability, and liquidity and asset qualities. Eleven ratios of the twelve core FSIs used in this study are:

- Regulatory Capital to Risk-Weighted Assets
- Regulatory Tier 1 Capital to Risk-Weighted Assets
- Nonperforming Loans Net of Provisions to Capital
- Nonperforming Loans to Total Gross Loans
- Return on Assets
- Return on Equity
- Interest Margin to Gross Income
- Noninterest Expenses to Gross Income
- Liquid Assets to Total Assets
- Liquid Assets to Short-Term Liabilities
- Net Open Position in Foreign Exchange to Capital

On the other hand, encouraged FSIs are the ratios which supply information regarding equity quality, credit quality and financial derivatives etc. There are also twelve encouraged FSIs and seven of them are used in this study:

- Capital to Assets
- Gross Asset and Liability Position in Financial Derivatives to Capital
- Trading Income to Total Income
- Personnel Expenses to Noninterest Expenses
- Customer Deposits to Total (Noninterbank) Loans
- Foreign-Currency-Denominated Loans to Total Loans
- Foreign-Currency-Denominated Liabilities to Total Liabilities

FSIs are calculated using the data available in financial statements which are published on websites of The Participation banks and The Banks Association of Turkey (TBB), for participation banks and banks respectively. Data in these financial tables are suitably transformed so that they will be in compliance with balance sheet and income statement forms proposed in FSI Compilation Guide (2006). Using these newly formed financial statements, FSIs have been calculated. Since public financial statements and the financial statements in compliance with FSI Compilation Guide have different forms, data corresponding to different accounts in financial statements needs to be calibrated. This is one of the main problems that has been faced during the data collection phase. Moreover, since the data for different banks is presented distinctly in different reports, these data had to be calibrated and transformed for each of 27 banks.

Some of the banks have been excluded from the analysis because of incomplete and too much deviated data. In that sense, banks under the control of Savings Deposit Insurance Fund (SDIF), foreign banks operating as branches in Turkey and development and investment banks are omitted from the study as well as Adabank. In other words, data related to the remaining 27 banks which are state-owned deposit banks, privately-owned deposit banks, foreign banks and four participation banks founded in Turkey have been analyzed. However, "sectoral distribution of loans to total loans", "large exposures to capital", "geographical distribution of loans to total loans", "spread between reference lending and deposit rates", "spread between highest and lowest interbank rates and "net open position in equities to capital" ratios are omitted from this analysis since the data required for calculation of these indicators cannot be obtained from *public sources*.

The readers may refer to FSI Compilation Guide (2006) for detailed information on calculation of FSIs and on other complementary information. In appendix, box plots of FSIs belonging to March 2012 are shown to represent general behavior of data set used in study.

4.2. Application

In this study we use fuzzy c-means clustering and FSIs to classify Turkish commercial banks regarding their soundness. As it is explained before, there are two main classes of data clustering; hard and fuzzy clustering. The main advantage of fuzzy clustering over the hard clustering is its probabilistic nature. Although some hard clustering methods such as K-means and especially K-medoids are more robust to noise and outliers then the other clustering techniques, they do not explain financial data as much as fuzzy clustering methods do. As it is explained in next paragraph they break down the continuum of assessment process. Furthermore, these techniques do not have the power to assess the data falling in between two sets. To eliminate outlier effects, this study concentrates on most similar banks, namely commercial banks, whose data are not too much deviated from the chosen sample. Although, the FSIs used in this study do not fully exclude the effects of noise and outliers, the convergence of FCM is not too much affected as shown in Table 3.

Fuzzy logic systems can be seen as a special case of classical expert systems which produce software solutions trying to recreate human decision-making ability in a specific area of application. In contrast to expert systems, specific values which are entered for credit quality of a firm can be assigned to multiple terms using various degrees of membership instead of being allocated to a single linguistic term (e.g. high: low) under fuzzy clustering methods. For instance, under a classical expert system, a criterion may be conditioned on the value of a certain financial ratio. As an example, this criterion may be defined such that the ratio having a value larger than or equal to 20% is attributed to be good or not good otherwise. It can be easily seen that this dual assignment is not in line with human assessment behavior. If 20% is attributed to be good then 19.90% is not bad either in accordance with the way human beings think. Fuzzy logic systems mostly resemble human decision-making behavior by introducing linguistic variables since they enable a tender gradation (OeNB & FMA, November 2004). Therefore, the models depending on fuzzy logic dominate classical expert systems and hard clustering algorithms regarding the clustering of financial data due to their resemblance to human decision making ability. Thus, in this study we use FCM, one of the fuzzy-logic based algorithms, to classify 27 of Turkish commercial banks.

After applying FCM clustering algorithm to FSIs of commercial banks which belong to March June and September 2012 periods, the results that are presented in Table 2, Table 4 and Table 5 are obtained respectively. For the purpose of better representation, the matrix in Equation (1) has been transposed to get each of those tables. For this representation, class number c and fuzziness parameter m are taken as 6 and 2 consecutively. Table 3 gives the results of FCM for different values of c and m regarding March 2012 data. Although objective function (FCM) attains its minimum at higher values of fuzziness coefficient m and higher values of class c, the probabilities of membership get too closer for each bank for that high m and c values. Considering FCM value for $c = 6$ and $m = 2$ we have better distinguished classes for each bank, because larger c and m values decrease the value of membership functions. The values of membership functions get closer and this makes difficult to draw clear boundaries between clusters. The low values for FCM make sense since number of banks studied is relatively small in this work and for chosen parameters c and m the values of FCM are sufficiently small.

When Table 2 is analyzed, from general subjective assessments one can propose that the results are really sensible. Because, according to the results, all participation banks fall in the same class (class 1), together with Şekerbank and Fibabanka. However, the degree of membership to class 1 is smaller for Şekerbank and Fibabanka when compared to participation banks. Şekerbank's and Fibabanka's membership function values for class 3 are also high. Due to high membership to both class 1 and class 3, this method with the given parameters fails to strongly reject that these two banks do not belong to either one of these clusters. To exemplify, Şekerbank's membership function takes value 28.65% for class 1 and 25.60% for class 3 whereas Fibabanka's membership function has values 21.86 and 20.10% for class 1 and class 3 respectively. This can be taken as a sign of great similarity between Şekerbank and Fibabanka.

From Table 2 we can conclude that Citibank is the basic representative of the class 2 with its highest membership function value (79.70%). As it is observed, Turkland Bank has second higher value (54.34%) for this class. Burgan[2] Bank has nearly same values for this class and class 4. This can be explained in such a way that Burgan Bank could be included in either of the classes 2 and 4, which indicates that this bank possesses financial stability characteristics of both of these classes.

Table 2: Degree of membership, March 2012

Bank Name	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6
ALBARAKA TÜRK KATILIM BANKASI	33,18%	6,75%	20,59%	7,37%	14,12%	17,98%
ASYA KATILIM BANKASI	40,50%	3,08%	26,05%	3,92%	15,43%	10,93%
FİBABANKA	21,86%	7,48%	20,10%	13,85%	18,07%	18,64%
KUVEYT TÜRK KATILIM BANKASI	40,12%	3,35%	22,89%	4,66%	18,14%	10,84%
ŞEKERBANK	28,65%	6,83%	25,26%	6,66%	13,96%	18,64%
TÜRKİYE FİNANS KATILIM BANKASI	46,69%	2,75%	19,54%	4,58%	12,23%	14,21%
CITIBANK	3,60%	79,70%	3,67%	4,11%	3,10%	5,81%
BURGAN BANK	14,02%	21,80%	13,63%	20,16%	11,11%	19,28%
TURKISH BANK	14,68%	29,01%	14,64%	10,29%	14,20%	17,19%
TURKLAND BANK	7,25%	54,34%	6,80%	12,08%	5,41%	14,10%
DENIZBANK	25,11%	2,47%	43,08%	4,05%	17,39%	7,90%
FİNANSBANK	23,30%	4,91%	34,57%	6,47%	15,31%	15,44%
TÜRK EKONOMİ BANK	19,72%	2,28%	55,98%	3,34%	10,71%	7,96%
YAPI VE KREDİ BANKASI	18,17%	2,18%	41,51%	3,25%	27,67%	7,20%
ALTERNATIFBANK	2,89%	2,42%	2,89%	84,73%	2,15%	4,91%
ARAP TÜRK BANKASI	14,94%	10,84%	15,95%	28,39%	14,07%	15,81%
DEUTSCHE BANK	13,17%	17,92%	14,47%	24,27%	14,02%	16,16%
AKBANK	12,15%	2,68%	18,63%	3,64%	54,45%	8,45%
TC. ZİRAAT BANKASI	18,53%	16,42%	16,07%	9,85%	19,74%	19,38%
TÜRKİYE GARANTİ BANKASI	14,96%	3,40%	17,51%	4,28%	51,45%	8,41%
TÜRKİYE HALK BANKASI	17,25%	4,64%	15,91%	5,30%	43,31%	13,59%
TÜRKİYE İŞ BANKASI	11,86%	1,70%	12,16%	2,45%	65,70%	6,13%
TÜRKİYE VAKIFLAR BANKASI	15,42%	1,94%	15,02%	2,68%	58,47%	6,47%
ANADOLUBANK	14,18%	7,96%	11,72%	16,01%	8,03%	42,10%
HSBC BANK	16,50%	14,11%	18,65%	10,40%	13,15%	27,18%
ING BANK	19,05%	6,53%	21,97%	11,91%	13,09%	27,45%
TEKSTİL BANKASI	8,28%	3,13%	6,42%	3,80%	4,84%	73,53%

[2] Eurobank Tekfen A.Ş. has been renamed as Burgan Bank A.Ş. as of 01.23.2013.

Table 3: Objective Function

Class \Fuzziness coefficient	m=2	m=3	m=4
c=4	1,13	0,29	0,09
c=5	0,90	0,19	0,04
c=6	0,74	0,13	0,02
c=7	0,63	0,09	0,01

The basic distinguishable property of class 3 is that all members of this class have quite high value of membership to this class compared to their membership degree to the other classes. Another feature of this class is also interesting. All members of this class has second highest membership degrees to the class 1. Therefore, according to FSIs (one of the widely used financial indicators class), it can be claimed that the members of this classes have comparably similar financial stability.

According to results of March 2012 data, all the publicly-owned banks, Ziraat Bank, Halk Bank and Vakıf Bank, are the members of the same class (class 5). Besides these public-owned banks, İşbank, Akbank and Garanti Bank are also members of that class. One of the basic properties of that class is high membership of all banks except from Ziraat Bank. That is, we can claim that all members of these classes represent same features apart from Ziraat Bank. As it can be observed, membership function of Ziraat Bank takes nearly same values for all classes except class 4.

Table 4: Degree of membership, June 2012

Bank Name	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6
ALTERNATIFBANK	23,43%	22,82%	11,03%	6,93%	12,78%	23,01%
CITIBANK	21,01%	23,06%	13,08%	10,19%	12,13%	20,54%
BURGAN BANK	25,35%	25,77%	10,98%	6,10%	7,55%	24,25%
TURKISH BANK	20,21%	21,78%	14,33%	11,70%	12,20%	19,79%
TURKLAND BANK	25,66%	30,13%	8,23%	4,85%	6,29%	24.S4%
ALBARAKA TÜRK KATILIM BANKASI	8,72%	8,15%	42,45%	13,50%	18,28%	8,90%
ASYA KATILIM BANKASI	7,29%	6,71%	37,02%	14,22%	27,30%	7,46%
FİBABANKA	18,16%	16,94%	18,87%	12,51%	15,33%	18,19%
KUVEYT TÜRK KATILIM BANKASI	8,05%	7,58%	47,91%	13,24%	15,09%	8,12%
ŞEKERBANK	16,09%	14,22%	25,27%	9,21%	18,64%	16,57%
TC. ZİRAAT BANKASI	15,65%	15,85%	19,71%	18,71%	14,36%	15,71%
TÜRKİYE FİNANS KATILIM BANKASI	8,08%	7,27%	49,97%	9,62%	16,73%	8,33%
AKBANK	3,84%	3,52%	7,56%	70,43%	10,71%	3,93%
TÜRKİYE GARANTİ BANKASI	3,92%	3,65%	8,42%	70,41%	9,61%	3,99%
TÜRKİYE HALK BANKASI	10,09%	9,67%	18,70%	37,03%	14,26%	10,24%
TÜRKİYE İŞ BANKASI	6,46%	5,96%	19,45%	47,94%	13,63%	6,57%
TÜRKİYE VAKIFLAR BANKASI	4,95%	4,49%	14,64%	53,05%	17,78%	5,09%
DENIZBANK	3,44%	3,05%	11,26%	8,66%	70,06%	3,54%
FİNANSBANK	8,06%	7,26%	16,81%	12,34%	47,21%	8,31%
ING BANK	10,99%	9,86%	15,21%	10,74%	41,89%	11,30%
TÜRK EKONOMİ BANKASI	5,00%	4,49%	13,18%	10,32%	61,88%	5,13%
YAPI VE KREDİ BANKASI	6,82%	5,97%	16,88%	23,65%	39,67%	7,02%
ANADOLUBANK	20,32%	18,91%	20,18%	8,19%	11,48%	20,92%
ARAP TÜRK BANKASI	18,77%	18,17%	14,85%	13,49%	15,90%	18,82%
DEUTSCHE BANK	18,78%	18,70%	13,72%	14,90%	15,09%	18,81%
HSBC BANK	19,32%	18,16%	13,39%	10,09%	19,05%	19,98%
TEKSTİL BANK	25,58%	23,66%	10,01%	5,60%	8,09%	27,06%

Comparing the results of June data with March data, participation banks, Şekerbank and Fibabanka are also fall in the same class. Furthermore, according to June data Ziraat Bank moves the same class as well. As it is mentioned before, this is not too much surprising, since according to March data, Ziraat Bank has nearly same value of membership to class composed of participation banks, Şekerbank and Fibabanka. Regarding the data of this term, without any exception Ziraat Bank takes nearly the same membership to all classes without any exception. Similar to preceding period, Citibank, Burgan Bank, Turkish Bank and Turkland Bank constitute same classes. One of the distinguishable property of these term results from the previous term is that Alternatifbank constitutes a single class. However, as it can be observed from

Table 2, this bank is a member of class formed by Arap Türk Bank and Deutche Bank. According to June data results Alternatifbank takes nearly the same value of membership for class containing these banks.

Table 5: Degree of membership, September 2012

Bank Name	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6
ALBARAKA TÜRK KATILIM BANKASI	22.46%	16.55%	6.02%	22.00%	14.45%	18.51%
FIBABANKA	21.83%	17.09%	3.85%	21.59%	17.42%	18.21%
KUVEYT TÜRK KATILIM BANKASI	20.97%	19.75%	9.33%	20.79%	12.94%	16.22%
ŞEKERBANK	24.43%	17.74%	3.08%	23.90%	11.36%	19.49%
TÜRKİYE FİNANS KATILIM BANKASI	24.11%	11.89%	2.43%	22.91%	14.74%	23.92%
ALTERNATIFBANK	19.48%	33.66%	6.98%	20.01%	7.98%	11.88%
ARAP TÜRK BANK	19.10%	20.11%	9.77%	19.20%	15.25%	16.57%
DEUTSCHE BANK	17.64%	20.05%	14.49%	17.82%	14.55%	15.45%
BURGAN BANK	20.21%	35.82%	5.49%	20.85%	7.93%	9.69%
TURKLAND BANK	17.24%	27.78%	19.47%	17.80%	7.56%	10.16%
CITIBANK	2.47%	2.89%	88.16%	2.49%	1.85%	2.13%
TC. ZİRAAT BANKASI	16.27%	15.42%	18.86%	16.24%	17.78%	15.44%
TURKISH BANK	11.64%	13.34%	44.77%	11.75%	8.57%	9.94%
ANADOLUBANK	24.61%	21.70%	4.49%	24.90%	10.43%	13.87%
TEKSTİL BANKASI	25.18%	21.64%	3.85%	25.92%	9.09%	14.32%
AKBANK	11.85%	7.44%	2.11%	11.46%	47.36%	19.78%
TÜRKİYE GARANTİ BANKASI	10.63%	7.56%	3.42%	10.34%	51.88%	16.16%
TÜRKİYE HALK BANKASI	12.57%	8.90%	3.96%	12.29%	46.54%	15.74%
TÜRKİYE İŞ BANKASI	9.15%	5.88%	1.99%	8.87%	61.56%	12.56%
TÜRKİYE VAKIFLAR BANKASI	10.88%	6.38%	1.98%	10.49%	49.01%	21.26%
ASYA KATILIM BANKASI	17.06%	9.32%	3.19%	16.25%	19.89%	34.29%
DENİZBANK	12.17%	6.89%	2.34%	11.68%	18.38%	48.54%
FİNANSBANK	16.52%	10.37%	3.29%	16.03%	17.25%	36.54%
HSBC BANK	21.60%	18.19%	5.77%	21.73%	10.44%	22.27%
INGBANK	19.14%	12.19%	3.19%	18.75%	11.77%	34.96%
TÜRK EKONOMİ BANK	13.58%	7.44%	2.05%	12.97%	12.05%	51.91%
YAPI VE KREDİ BANKASI	16.80%	9.45%	2.05%	16.06%	23.60%	32.03%

The results of September 2012 data are also not very different from results of other two periods either. One of the main differences from other two periods is that one of the participation banks, Asya Katılım Bank, falls into another class. Also Ziraat Bank falls into the class including Citibank and Turkish bank.

Similar to preceding two periods, in this period Akbank, Garanti Bank, İş Bank, Halk Bank and Vakıf Bank constitute same class. Furthermore, Denizbank, Finansbank, Ekonomi Bank and Yapı Kredi Bank remain in the same class. Ziraat Bank moves to other class that is formed by Citibank and Turkish Bank.

5. Conclusion

Banks are the main actors of financial markets. Thus, financial stability of market is very rigidly related with soundness of banks. The aim of this paper is classifying 27 of Turkish commercial banks (deposit and participation banks) regarding banks' financial soundness via FSIs. Although there are several data classification methods, fuzzy c-means algorithm (FCM) is used for classification purposes. The main reason for choosing FCM is its similarity to human decision-making behavior. This property of FCM comes from its probabilistic nature.

After applying FCM to three periods, March, June and September 2012, we get separated classes of banks with respect to membership functions for each period. According to membership degrees of all banks considered in all three periods, one obvious inference is that nearly in all periods, participation banks fall in the same cluster, with the exception of Asya Katılım Bank, which moves into another class according to September 2012 data.

Another conspicuous property of this work is that, pursuant to results obtained using data for each of the three periods; Denizbank, Finansbank, Yapı Kredi Bank and Ekonomi Bank fall in the same cluster. Also, Akbank, Garanti Bank, İş Bank, Halk Bank and Vakıf Bank are members of the same classes according to three period results. These results are very astonishing in accordance with the general perception about financial conditions of these banks. First of all, considering Denizbank, Finansbank, Yapı Kredi Bank and Ekonomi Bank, all these banks are private commercial banks and each bank has foreign joint. Considering their credit card services, derivative position and asset sizes (apart from Yapı Kredi Bank) intuitively we can claim that there exists a great similarity among these banks. For instance, when Akbank, Garanti Bank, İş Bank, Halk Bank and Vakıf Bank are considered, it can be safely said that these banks are leading banks of Turkish banking sector with Ziraat Bank and Yapı Kredi Bank.

As it is concluded from the results, asset size is not the dominating variable which heavily affects the results of this study. Yapı Kredi Bank which is the fifth largest bank with respect to its asset size according to March 2012 data does not fall in the same class with other largest banks, İş Bank, Ziraat Bank, Garanti Bank, Akbank, Vakıfbank and Halk Bank. This fact shows that the result of study is not affected by just one observation.

Since fuzzy c-means algorithm is a classification tool, just by using the result of this algorithm we cannot decide which class is superior. But one can use some other judgment tools to decide which class has better credit quality. Therefore, the aim of this study is not to rate or sort banks according to their financial soundness. Nevertheless, the aim of this work is to assess the resemblance of Turkish commercial banks via FSIs. However, it can be argued that the results of this work supply important clues regarding the financial structure of the banks studied in this work.

The basic weakness of this study results from the absence of model calibration. The reason behind this absence is the shortage of data. However, this work has a great importance because of being the first study that evaluates soundness of Turkish banking sector by using one of the globally accepted clustering models.

References

1. Alam, P., Booth, D., Lee, K., & Thordarson, T. (2000). The use of fuzzy clustering algorithm and self-organizing neural networks for identifying potentially failing banks: an experimental study. *Expert Systems with Applications*, 185-199.
2. Babuska, R. (2009). *Fuzzy and neural control disc: Course lecture notes.* Delft, the Netherlands: Delft University of Technology.
3. Bellman, R., Kalaba, R., & Zadeh, L. (1996). Fuzzy sets, fuzzy logic, and fuzzy systems. *Abstraction and pattern classification*, 44-50.
4. Bezdek, J. C. (1981). *Pattern Recognition with Fuzzy Objective Function Algorithms.* Norwell, MA. USA: Kluwer Academic Publishers.
5. Bezdek, J. C., Ehrlich, R., & Full, W. (1984). FCM: The fuzzy c-means clustering algorithm. *Computers & Geosciences 10(2-3)*, 191-203.
6. Chen, L.-H., & Chiou, T.-W. (1999). A Fuzzy Credit-rating Approach for Commercial Loans: a Taiwan Case. *Omega, Int. J. Mgmt. Sci.*, 407-419.
7. Connor, G., Flavin, T., & O'Kelly, B. (2012). The U.S. and Irish credit crises: Their distinctive differences and common features. *Journal of International Money anc Finance*, 60-79.
8. Doğan, B. (2008). *CLUSTER ANALYSIS AS A BANKING SUPERVISION TOOL: AN APPLICATION FOR TURKISH BANKING SECTOR.* İstanbul: Kadir Has University.
9. Dunn, J. C. (1973). A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters. *Mathematical and Computer Modelling*, 32–57.
10. Gökgöz, İ. H. (2012). Stochastic credit default swap pricing.
11. International Monetary Fund. (2006). *Financial Soundeness Indicators.* Washington, D.C.: International Monetary Fund Publication Services .
12. Krishnapuram, R. (1993). A possibilistic approach to clustering. *IEEE Transactions on Fuzzy Systems*, 98-110.

13. OeNB, & FMA. (November 2004). *Guidelines on credit risk management: Rating models and validation.* Vienna, Austria: OeNB Printing Office.

14. Pedrycz, W. (2005). *Knowledge-Based Clustering: From Data to Information Granules.* New Jersey: John Wiley & Sons. Inc.

15. Ruspini, E. H. (1969). A new approach to clustering. *Information and Control*, 22-32.

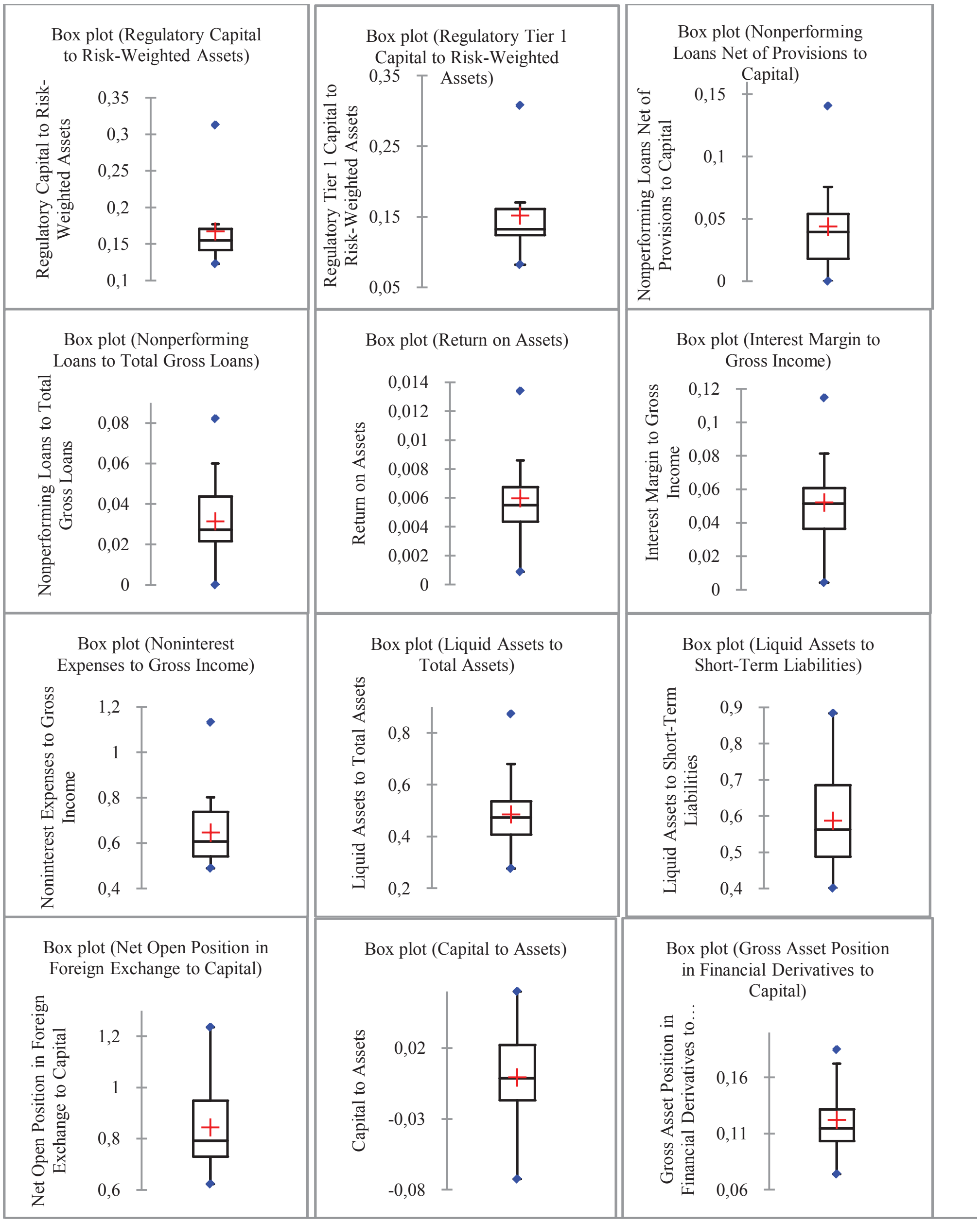
16. Servigny, A. D., & Renault, O. (2004). *Measuring and Managing Credit Risk.* New York: The McGraw-Hill Companies. Inc.

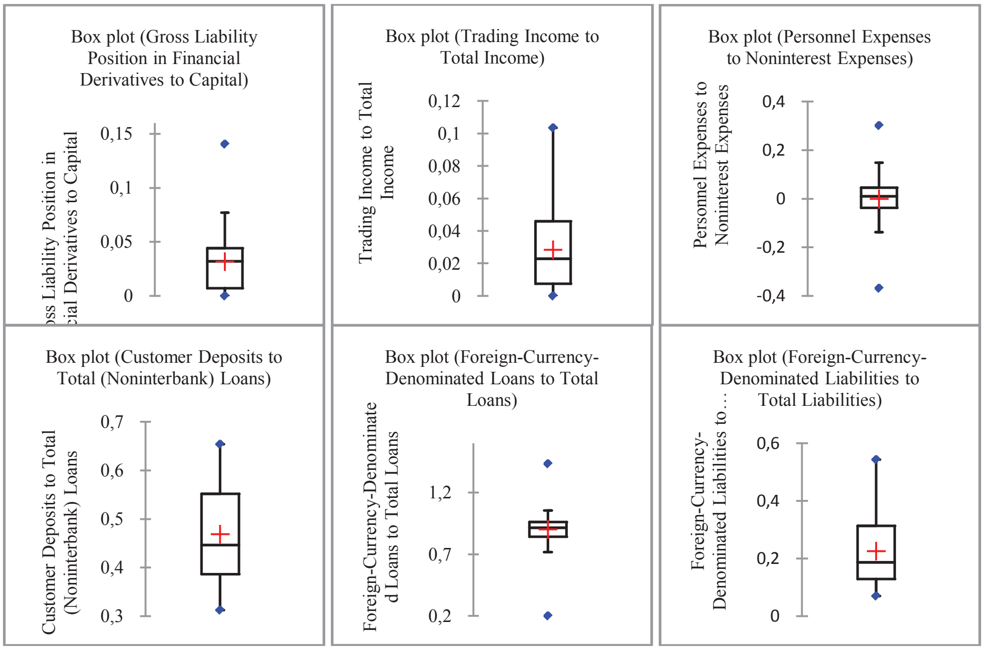
17. Tufan, E., & Hamarat, B. (2003). Clustering of Financial Ratios of the Quoted Companies Through Fuzzy Logic Method. *Journal of Naval Science and Engineering*, 123-140.

18. Yang, M. S. (1993). A survey of fuzzy clustering. *Mathematical and Computer Modelling 1S(11)*, 1-16.

19. Zadeh, L. A. (1965). Fuzzy sets. *Information and Control Vol 8*, 338-353.

21. Appendix: March Data Box Plots




Box plot (Gross Liability Position in Financial Derivatives to Capital)
0,15
0,1
0,05
0
Box plot (Trading Income to Total Income)
Trading Income to Total Income
0,12
0,1
0,08
0,06
0,04
0,02
0
Box plot (Personnel Expenses to Noninterest Expenses)
Personnel Expenses to Noninterest Expenses
0,4
0,2
0
-0,2
-0,4
Box plot (Customer Deposits to Total (Noninterbank) Loans)
Customer Deposits to Total (Noninterbank) Loans
0,7
0,6
0,5
0,4
0,3
Box plot (Foreign-Currency-Denominated Loans to Total Loans)
Foreign-Currency-Denominated Loans to Total Loans
1,2
0,7
0,2
Box plot (Foreign-Currency-Denominated Liabilities to Total Liabilities)
Foreign-Currency-Denominated Liabilities to…
0,6
0,4
0,2
0