Öner / Journal of Economics and Administrative Sciences, 27(2): 461-489, 2026
DOI: 10.37880/cumuiibf.1885630

Journal of Economics and Administrative Sciences

| esjournal.cumhuriyet.edu.tr | Founded: 2000 Available online, ISSN: 1303-1279 Publisher: Sivas Cumhuriyet Üniversitesi

CLASSIFICATION OF PROVINCES USING CLUSTERING METHODS IN TERMS OF BANK DEPOSITS AND LOANS

Burçin Öner[1*]

[1*] Correspondence Author, Bilecik Şeyh Edebali University, Bilecik, Türkiye, burcin.oner@bilecik.edu.tr; **ORCID:** 0000-0001-9550-0435

Abstract: Numerous studies in the literature classify provinces in Turkey using various indicators such as socio-economic development, education, and health, and examine clustering structures based on similarities or differences. This study differs by classifying provinces according to bank deposits and bank loans and evaluating their clustering patterns. Six variables related to deposits and loans were used, and the data were obtained from the 2022 statistics of the Turkish Banks Association. Per capita bank deposits and loans were calculated using 2022 provincial population figures, and clustering analyses were conducted based on these per capita values. As the number of clusters, six tiers defined in the SEGE2017 socio-economic development rankings of the Ministry of Industry and Technology were adopted. Additionally, changes in clustering structures were examined for different cluster counts. Although various classification studies exist, this study represents a new attempt in terms of socio-economic development levels, as it evaluates provinces according to SEGE2017 while using financial indicators. For clustering, k-means, complete linkage, Ward's method, and median clustering were applied. Across all methods, results based on both per capita deposits and loans exhibit notable similarities regardless of the number of clusters. Istanbul and Ankara consistently appear together in a single cluster, while Izmir, Antalya, Bursa, and Mugla are grouped in another. This may be due to the concentration of population in these provinces and their higher level of development compared to other provinces. Most provinces fall into two broad categories and are therefore not compatible with the SEGE2017 provincial classification structure.

Keywords: Classification, Clustering analysis, Deposit, Loan

JEL COdes: C19, C49

BANKALARDAKİ MEVDUAT VE KREDİLER AÇISINDAN KÜMELEME YÖNTEMLERİ KULLANILARAK İLLERİN SINIFLANDIRILMASI

Özet: Literatürdeki birçok çalışma, Türkiye'deki illeri sosyoekonomik gelişmişlik, eğitim ve sağlık gibi çeşitli göstergeler kullanarak sınıflandırmakta ve benzerlikler veya farklılıklara dayalı kümelenme yapılarını incelemektedir. Bu çalışma, illeri banka mevduatları ve banka kredileri açısından sınıflandırarak kümelenme modellerini değerlendirmesi bakımından farklılık göstermektedir. Mevduat ve kredilerle ilgili altı değişken kullanılmış ve veriler Türkiye Bankalar Birliği'nin 2022 istatistiklerinden elde edilmiştir. Kişi başına banka mevduatları ve kredileri, 2022 il nüfus rakamları kullanılarak hesaplanmış ve bu kişi başına değerlere dayalı olarak kümelenme analizleri yapılmıştır. Küme sayısı olarak, Sanayi ve Teknoloji Bakanlığı'nın SEGE2017 sosyo-ekonomik gelişmişlik sıralamasında tanımlanan altı kademe benimsenmiştir. Ek olarak, farklı küme sayıları için kümelenme yapılarındaki değişiklikler incelenmiştir. Çeşitli sınıflandırma çalışmaları mevcut olmakla birlikte, bu çalışma, finansal göstergeler kullanılarak illeri SEGE2017'ye göre değerlendiren, sosyoekonomik gelişmişlik düzeyleri açısından yeni bir girişimdir. Kümelenme için k-means, tam bağlantı, Ward yöntemi ve medyan kümelenme uygulanmıştır. Tüm yöntemlerde, kişi başına mevduat ve kredilere dayalı sonuçlar, küme sayısına bakılmaksızın dikkate değer benzerlikler göstermektedir. İstanbul ve Ankara sürekli olarak tek bir kümede yer alırken, İzmir, Antalya, Bursa ve Muğla başka bir kümede gruplandırılmıştır. Bunun nedeni olarak nüfus yoğunluğunun bu illerde toplanmış ve bu illerin gelişmişlik düzeyinin diğer illere göre daha yüksek olması olabilir. Çoğu il, iki geniş kümeye girmiş ve bu yönden SEGE2017 il sınıflandırma yapısı ile uyumluluk göstermemektedir.

Keywords: Sınıflandırma, Kümeleme analizi, Mevduat, Kredi

JEL Kodu: C19, C49

Introduction

In the banking sector, deposits and loans are considered two vital operational activities for a bank. A deposit refers to money entrusted to institutions authorized to collect funds, either remunerated (in exchange for a fee, right, or claim) or non-remunerated (without compensation), and with the condition of repayment (Şener, 2019; Er, 2022). A loan, on the other hand, is a concept with various definitions; however, its most common usage refers to the provision of a certain amount of purchasing power to an individual or legal entity, with a promise of repayment at the end of a specified term, in exchange for a fee (interest) (Köklü, 1967; Çağıran ve Ulusoy, 2019). It is possible to state that the deposit process, which is of vital importance for banks, also makes a significant contribution to the national economy. In this respect, it is a fact that deposits hold a crucial place among the resources of banks. Banks collect deposits at relatively low interest rates and extend them as loans to those in need at higher interest rates through loan agreements. Consequently, the interest rate difference constitutes the profit of the bank(s).

Any differences observed between countries or between regions or provinces within a country are characterised as regional imbalances. It is possible to measure and evaluate the development levels of provinces in a country based on various factors. These factors can be categorized as economic, social, educational, historical, geographical, climatic, financial, strategic, etc. (Karabulut vd., 2004). Like in many other fields, Turkey has undertaken efforts in the statistical domain to fulfill the obligations of European Union (EU) member states. In this context, the Turkish Statistical Institute (TUIK) and the State Planning Organization (DPT) established the Statistical Regional Units Classification (SRUC) in 2001. This classification was published in the Legal Gazette on September 22, 2022, following the decision of the Council of Ministers (2002/4720). Two main criteria were considered in the preparation of SRUCs. The first of these criteria requires that each higher level be formed by combining its own subordinate units, and the second requires that regional units form a continuous whole with each other in terms of geographical boundaries (Erilli, 2014b; URL-1). The second criterion is ensuring that regional units maintain geographical continuity and integrity.

The SRUC consists of three levels. According to the classification, Level-1 comprises 12 statistical regional units, Level-2 consists of 26 regions, and Level-3 represents 81 provinces. After designating provinces as Level-3, Level-2 regions were determined.

In Turkey, provinces were classified as Level-3 due to their administrative structure and the collection of statistical data (e.g., population, Ağrı (04) culture, industry) primarily at the provincial level, aligning with the formation of NUTS-3 units in EU member states. Additionally, as NUTS-2 units in EU member states are the regions that receive the most aid, this criterion was also taken into account when defining Level-2 regions. Accordingly, Level-2 regions were created from provinces that are socioeconomically and culturally similar, geographically comparable, and share common challenges. The same approach was applied to the formation of Level-1 regions. Table 1 presents the Level-1 regions, while Table 2 provides the Level-2 regions and their regional center provinces.

Table 1. Level-1 regions

Code	Level-1 Region	Code	Level-1 Region
TR1	İstanbul	TR7	Centeral Anatolia
TR2	Western Marmara	TR8	Western Black Sea
TR3	Aegean	TR9	Eastern Black Sea
TR4	Eastern Marmara	TRA	Northeastern Anatolia
TR5	Western Anatolia	TRB	Middleeastern Anatolia
TR6	Mediterranean	TRC	Southeastern Anatolia

Table 2. Level-2 regions

Code	Level-2 Region	Code	Level-2 Region
TR10	İstanbul (34)	TR71	Kırıkkale (71)
TR21	Tekirdağ (59)	TR72	Kayseri (38)
TR22	Balıkesir (10)	TR81	Zonguldak (67)
TR31	İzmir (35)	TR82	Kastamonu (37)
TR32	Aydın (09)	TR83	Samsun (55)
TR33	Manisa (45)	TR90	Trabzon (61)
TR41	Bursa (16)	TRA1	Erzurum (25)
TR42	Kocaeli (41)	TRA2	Ağrı (04)
TR51	Ankara (06)	TRB1	Malatya (44)
TR52	Konya (42)	TRB2	Van (65)
TR61	Antalya (07)	TRC1	Gaziantep (27)
TR62	Adana (01)	TRC2	Şanlıurfa (63)
TR63	Hatay (31)	TRC3	Mardin (47)

As an example of the SRUC, which facilitates the collection, development, and harmonization of regional statistics while enabling socioeconomic analyses, Table 3 presents the Level-1, Level-2, and Level-3 information for the Southeastern Anatolia Region.

Table 3. Southeastern Anatolia NUTS

Code	Level-1 Region	Code	Level-2 Region	Code	Level-3 Region
TRC	Southeastern Anatolia	TRC1	Gaziantep (27)	TRC11	Gaziantep (27)
				TRC12	Adıyaman (02)
				TRC13	Kilis (79)
		TRC2	Şanlıurfa (63)	TRC21	Şanlıurfa (63)
				TRC22	Diyarbakır (21)
		TRC3	Mardin (47)	TRC31	Mardin (47)
				TRC32	Batman (72)
				TRC33	Şırnak (73)
				TRC34	Siirt (56)

The objectives of establishing SRUCs can be summarized as follows: collecting and developing regional statistics, conducting socioeconomic analyses of regions, defining the framework for regional policies, and presenting fundamental indicators that contribute to the establishment of a comparable statistical database aligned with the EU Regional Statistics System (Erilli, 2014a). Moreover, in the SEGE 2017 report prepared by the General Directorate of Development Agencies under the Ministry of Industry and Technology, provinces were classified into six tiers based on their socioeconomic development rankings (SEGE, 2017). The principal component values obtained through principal component analysis were used to establish these rankings. The distribution of provinces across the six tiers is shown in Table 4.

In the literature, numerous studies classify Turkish provinces based on various indicators such as socio-economic development, educational services, and health indicators, analyzing their clustering structures according to similarities and differences. Unlike previous studies in the literature, this study aims to classify provinces and evaluate their clustering structures based on per capita bank deposits and bank loans.

Table 4. Distribution of provinces by SEGE 2017 rankings (SEGE, 2017)

Tier	Number of provinces	Provinces
1	9	İstanbul (34), Ankara (06), İzmir (35), Kocaeli (41), Antalya (07), Bursa (16), Eskişehir (26), Muğla (48), Tekirdağ (59).
2	15	Denizli (20), Sakarya (54), Yalova (77), Bolu (14), Konya (42), Aydın (09), Isparta (32), Kayseri (38), Kırklareli (39), Bilecik (11), Çanakkale (17), Edirne (22), Karabük (78), Manisa (45), Balıkesir (10).
3	13	Mersin (İçel) (33), Trabzon (61), Adana (01), Zonguldak (67), Uşak (64), Gaziantep (27), Samsun (55), Burdur (15), Kırıkkale (71), Düzce (81), Karaman (70), Rize (53), Kütahya (43).
4	14	Amasya (05), Hatay (31), Nevşehir (50), Afyonkarahisar (03) (03), Elâzığ (23), Kırşehir (40). Malatya (44), Sivas (58), Bartın (74), Erzincan (24), Kastamonu (37), Artvin (08), Çorum (19), Aksaray (68).
5	14	Sinop (57), Giresun (28), Osmaniye (80), Çankırı (18), Tokat (60), Niğde (51), Kahramanmaraş (46), Tunceli (62), Ordu (52), Erzurum (25), Kilis (79), Yozgat (66), Gümüşhane (29), Bayburt (69).
6	16	Adıyaman (02), Ardahan (75), Diyarbakır (21), Kars (36), Iğdır (76), Bingöl (12), Batman (72), Şanlıurfa (63), Mardin (47), Siirt (56), Bitlis (13), Van (65), Hakkâri (30), Muş (49), Ağrı (04), Şırnak (73).

Importance of the Subject

The study assumes that population size may be a significant factor in the distribution of bank deposits and loans across provinces. Therefore, per capita bank deposits and per capita bank loans are used as the basis for analysis. By weighting deposits and loans -expected to be naturally higher in densely populated provinces-according to population size, an attempt is made to establish balance among provinces. Given the significance of banking transactions, including deposits and loans, for both the national economy and individual livelihoods, revealing the similarities among Turkish provinces in terms of per capita deposits and loans may also serve as an indicator of their level of development.

Literature Review

The process of grouping statistical units with some characteristics into the same category based on an evaluation of their similarities or differences is known as a classification or clustering problem. The statistical units in question may be living entities or inanimate objects. The literature contains studies on socio-economic data where provinces are treated as statistical units, as well as studies evaluating bank activities where banks are treated as statistical units. These studies focus on the classification of provinces, districts, regions, or banks.

In one of these studies, hierarchical clustering techniques were used to analyse the socio-economic similarities between provinces in Turkey, and the similarities between the results obtained from cluster analysis and those from multidimensional scaling analysis were discussed (Karabulut vd., 2004). In another study, the similarities between provinces in terms of socio-economic indicators were examined using cluster analysis and multidimensional scaling analysis (Kılıç vd., 2011). In another one, provinces in Turkey were classified according to their similarities in terms of banking activities using cluster analysis, and this classification analysis showed that the provinces were divided into six clusters (Yılmaz ve Uzgoren, 2013). In a similar study, clustering analysis was again used to compare indicators of access to and use of banking services across provinces (Sarıgül, 2014). A study that applied fuzzy clustering analysis to classify districts within the TR72 region according to socio-economic data showed that the districts were grouped into four clusters (Erilli, 2014b). In another study also fuzzy set analysis

was used to classify Level-2 regions and Level-3 regions, i.e. IBBSs, according to their level of development based on financial variables (Erilli, 2014a).

Studies have also been conducted in which the most important determinants of the socio-economic levels of provinces in Turkey were obtained through factor analysis, and the socio-economic development levels of provinces in 2012 were obtained through principal component analysis. This has provided a different perspective on determining the development levels of units (Albayrak ve Savaş, 2015). In a study on the classification of provinces using socio-economic data, researchers employed the Grey clustering method. In the first stage, grey correlation coefficient matrices were created using data obtained from the Turkish Statistical Institute (TÜİK). In the second stage, cluster analysis was used on variables identified in the subcategories of demographics, education, Agriculture and livestock, socio-economics, health, finance, and population and migration to determine the cluster structures of provinces separately in each category (Karadaş ve Erilli, 2023).

In a study that classified Turkish provinces based on banking data, credit usage levels were taken into account, and a dataset related to the types of credit financing provided by banks was utilized. Using this dataset, a clustering model for provinces was developed using the K-means method (Tekin ve Temelli, 2020a). Another study evaluating Turkish provinces in terms of banking activities used multivariate statistical analyses to examine similarities and differences among provinces based on selected variables (Yaman Yılmaz, 2020). This study applied principal component analysis (PCA) in the first stage to extract principal components from banking data variables. In the second stage, clustering analysis was conducted using both the original data and the principal component data to evaluate the clustering structures of provinces. In addition to studies that analyze the clustering structures of provinces using different variables related to socio-economic and banking factors, there are also studies that assess the clustering or ranking of banks based on their operational characteristics. These studies typically employ either multivariate statistical analyses or data envelopment analysis. Studies such as the classification of commercial banks in Turkey using financial ratios with a clustering analysis approach (Akgöz, 2010; Ayrıçay ve Akgöz, 2014) can be cited as examples of this (Önder ve Özyıldırım 2010) investigated the relationship between economic growth at the provincial level and public and private bank credits in Turkey using a fixed-effects model. Other notable studies in this field include those by Aydın ve Başkır (2013), Karaduman (2016), Meydan (2016), Bulut (2019), Tekin ve Temelli (2020b), Karaatlı ve Yıldız (2021), and Açıkalın (2023).

The primary distinction of our study from these previous works is that we analyze the clustering structures of provinces based on per capita bank deposits (using six deposit-related variables) and per capita bank loans (using six credit-related variables). This characteristic makes our study a contemporary contribution to the literature.

Material and Method

Clustering Analysis

Clustering analysis is one of the multivariate statistical analysis techniques. A cluster is a group formed by units that are similar or close to each other in certain characteristics, and the analysis applied to uncover such groups is called clustering analysis. It is one of the most widely used methods in classification studies. Cluster analysis, commonly used in the classification of units, is a method that classifies units within the research subject by grouping them into specific clusters based on their similarities or proximity, revealing their common characteristics and enabling general descriptions of these clusters (Tatlıdil, 2002).

Clustering analysis aims to partition units whose natural groups (classes) are not explicitly known into a limited number of groups or clusters. This partitioning process ensures that the units within the same cluster are similar to each other, while units in different clusters are distinct from one another. As a result of this process, intra-cluster homogeneity and inter-cluster heterogeneity emerge (Kalaycı, 2005; Alpar, 2020).

Clustering analysis became widely used after the 1960s and includes various clustering methods (Anderberg, 1973). The objectives of clustering analysis include identifying the true types to which samples belong, facilitating modeling, making preliminary estimates within groups, testing predefined hypotheses, clarifying data structure, reducing dimensionality, and detecting outliers (Tatlıdil, 2002; Alpar, 2020).

Clustering analysis is a collection of methods that help partition units, variables, or both into subgroups (clusters or classes) based on the data matrix of a research subject where natural groupings are not explicitly known (Özdamar, 2013). Clustering analysis methods use distance measures to assess similarities between units, where the closest units in terms of distance are considered the most similar.

In this study, hierarchical clustering methods -complete linkage method, median clustering method, and Ward's clustering method- along with the non-hierarchical K-means clustering method, were employed for unit clustering. Therefore, a brief introduction to these methods is provided below. In clustering analysis, a data matrix $\boldsymbol{X} = [x_{ji}]: p \times n; (j = 1,2, \dots, p$ and $i = 1,2, \dots, n)$ is used in a study with p variables and n sample units.

The distances between units are calculated using Euclidean distance or squared Euclidean distance. Let $\boldsymbol{x}_i' = [x_{1i}, \cdots, x_{pi}]$ and $\boldsymbol{x}_l' = [x_{1l}, \cdots, x_{pl}]$ represent the observation vectors of the i-th and l-th units, respectively. The Euclidean distance between these two observation vectors is calculated as follows:

$$d(\boldsymbol{x}_i, \boldsymbol{x}_l) = \sqrt{\sum_{J=1}^{p} (x_{ji} - x_{jl})^2} = \sqrt{(\boldsymbol{x}_i - \boldsymbol{x}_l)'(\boldsymbol{x}_i - \boldsymbol{x}_l)} \qquad (1)$$

Meanwhile, the squared Euclidean distance is given by:

$$d^*(\boldsymbol{x}_i, \boldsymbol{x}_l) = \sum_{J=1}^{p} (x_{ji} - x_{jl})^2 = (\boldsymbol{x}_i - \boldsymbol{x}_l)'(\boldsymbol{x}_i - \boldsymbol{x}_l) \qquad (2)$$

Thus, the matrix of Euclidean distances between units is denoted as $\boldsymbol{D} = [d_{il}]: n \times n$ and the matrix of squared Euclidean distances is denoted as $\boldsymbol{D}^* = [d_{il}^*]: n \times n$. In this study, squared Euclidean distance was used for Ward's clustering method, while Euclidean distance was used for the other methods.

Complete Linkage Clustering Method

The complete linkage method, also known as the farthest neighbor method, is a hierarchical clustering method. In this method, each unit is initially considered a separate cluster. First, a squared Euclidean distance matrix is created for the n units. In the initial stage, the two units with the smallest distance in the distance matrix are identified and merged to form a cluster. Subsequently, when determining the distances between the newly formed cluster and the other clusters, the clusters with the largest distance are identified. The algorithm for this method can be outlined as follows (Bulut, 2018):

i) The process starts by considering n units as n separate clusters, and the distance matrix D is calculated.

ii) The two closest clusters (i.e., the ones with the smallest $d(\boldsymbol{x}_i, \boldsymbol{x}_l)$) are merged.

iii) The number of clusters is reduced by one, and the distance matrix is updated. The updated distance between any clusters A and B is determined by taking the maximum distance among all units in those clusters:

$$d_{AB} = d(A, B) = max\{d(\boldsymbol{x}_i, \boldsymbol{x}_l)\} \quad (3)$$

iv) Steps (ii) and (iii) are repeated $n - 1$ times until all units are merged into a single cluster, or until a specific criterion is met. For example, the average distance between all units;

$$\overline{d} = \frac{2}{n(n-1)} \sum_{i<l=1}^{n} d_{il} \quad (4)$$

If $d_{il} > \overline{d}$ for all elements of the updated distance matrix, clustering is stopped.

Median Clustering Method

This method also uses the Euclidean distance matrix for clustering units. It is a hierarchical clustering method. The algorithm of the method can be described as follows.

i) The process begins by assuming n units and n clusters, and the D distance matrix is calculated.

ii) The two closest clusters (those with the smallest ($d(\boldsymbol{x}_i, \boldsymbol{x}_l)$) value) are merged. The cluster centre of these two merged clusters is calculated. The cluster centre of this new cluster $A = (\boldsymbol{x}_i, \boldsymbol{x}_l)$ is found as:

$$\overline{\boldsymbol{x}}_A = \frac{1}{2}(\boldsymbol{x}_i + \boldsymbol{x}_l) \quad (5)$$

iii) The number of clusters is reduced by one and the distance matrix is updated. In updating the distance matrix, the distance between any sets A and B is taken as the Euclidean distance between the centres of the two sets. The centre of a set is the average vector of the units in that set. Let $\overline{x}_A$ and $\overline{x}_B$ be the centres of sets A and B, respectively. The distance between these two sets is

$$d_{AB} = d(\overline{\boldsymbol{x}}_A, \overline{\boldsymbol{x}}_B) = \sqrt{(\overline{\boldsymbol{x}}_A - \overline{\boldsymbol{x}}_B)'(\overline{\boldsymbol{x}}_A - \overline{\boldsymbol{x}}_B)} \quad (6)$$

This distance is known as the Euclidean distance between cluster centres. If the smallest distance in the updated distance matrix is between sets A and B, these two sets are merged and the centre of the new set is calculated based on the median/midpoint. In this case, the new set centre will be:

$$m_{AB} = \frac{1}{2}(\overline{\boldsymbol{x}}_A + \overline{\boldsymbol{x}}_B) \quad (7)$$

iv) Steps (ii) and (iii) are repeated n-1 times until all units are grouped into a single cluster.

Ward's Clustering Method

This method is also known as Ward's least variance method. As a hierarchical clustering method, it is known to have a tendency to merge clusters with few observations and to form clusters with an equal number of units (Alpar, 2020). The Ward method uses the squared Euclidean distances defined by Equation (2) for similarities between units or clusters. The Ward

method aims to minimise the intra-cluster squared Euclidean distances and maximise the inter-cluster squared Euclidean distances. The sums of squares for clusters A and B are respectively;

$$KT_A = \sum_{i=1}^{n_A}(\boldsymbol{x}_i - \overline{\boldsymbol{x}}_A)'(\boldsymbol{x}_i - \overline{\boldsymbol{x}}_A) \quad (8)$$

$$KT_B = \sum_{i=1}^{n_B}(\boldsymbol{x}_i - \overline{\boldsymbol{x}}_B)'(\boldsymbol{x}_i - \overline{\boldsymbol{x}}_B) \quad (9)$$

These are expressed by the equations. If sets A and B are combined, the sum of squares within the new (AB) set is found using the following equation:

$$KT_{AB} = \sum_{k=1}^{n_{AB}}(\boldsymbol{x}_k - \overline{\boldsymbol{x}}_{AB})'(\boldsymbol{x}_k - \overline{\boldsymbol{x}}_{AB}) \quad (10)$$

Here, $n_{AB} = n_A + n_B$, while the centre of the (AB) set is calculated as:

$$\overline{\boldsymbol{x}}_{AB} = \frac{n_A\overline{\boldsymbol{x}}_A + n_B\overline{\boldsymbol{x}}_B}{n_{AB}} \quad (11)$$

and is known as the weighted average of sets A and B. The Ward method considers the increase in the sum of squares for the combined sets to be as small as possible. For this, the I_{AB} index is calculated for each case in all possible combinations using the formula:

$$I_{AB} = KT_{AB} - (KT_A + KT_B) \quad (12)$$

The sets with the minimum I_{AB} values are merged.

K-Means Clustering Method

It is a non-hierarchical clustering method and the most commonly used clustering method. There are two situations that distinguish this method from hierarchical clustering methods.

i) In hierarchical methods, the number of clusters is not known beforehand, but in the K-means clustering method, the number of clusters is known beforehand.

ii) Hierarchical methods require a distance matrix between units, but the K-means clustering method does not require a distance matrix. It is sufficient to know the distances of the units from the cluster centres.

This method aims to minimise intra-cluster variation and maximise inter-cluster variation. In other words, it aims to maximise intra-group homogeneity and inter-group heterogeneity. To achieve this goal, k initial values (cluster centres) called nuclei are first determined, and the analysis continues by assigning each observation to these nuclei in each iteration, thereby updating the nuclei. The objective function to be optimised in this method is given as following (Bulut, 2018):

$$\underset{m_1,\ldots,m_k}{\operatorname{argmin}} \sum_{i=1}^{n} \min_{j=1,2,\ldots,k} \left\|\boldsymbol{x}_i - \boldsymbol{m}_j\right\|^2 \quad (13)$$

A suitable algorithm for optimising the objective function given by Equation (13) can be provided as follows.

i. Initial cluster centres are determined (k number). The options that can be preferred in determining these core points are:

a) Randomly determine k core points

b) Take the first k observations as core points

c) Take the k observations that are furthest apart as core points

d) Take the cluster centres of k clusters obtained by the hierarchical clustering method as core points

ii. After k core points are determined, the remaining $n-k$ observations are assigned to the cluster containing the closest core point, and after each assignment, that cluster centre is recalculated and updated to be the average vector of the observations in the cluster at the time of assignment.

iii. The cluster centres of the k clusters obtained after step (ii) are accepted as core points. The n observations are reassigned to the nearest core cluster according to these core clusters. During this assignment process, some observations may be subject to transitions between clusters. The cluster centres are calculated for the clustering structure obtained at the end of the assignment process.

iv. Steps (ii) and (iii) are repeated until the transitions between clusters stabilise (Aggarwal and Reddy, 2014).

Creation of The Sample (Data Set)

Unlike studies in the literature, this study will classify provinces according to bank deposits and bank loans and evaluate their clustering structures. For this purpose, six variables related to bank deposits and bank loans were considered. These variables are presented in Table 5. The data for these variables were obtained and compiled from the 2022 data of the Turkish Banks Association (URL-2).

Table 5. Variables used in the research

Type	Variable	New variable $(Y_j = X_j/N)$
Deposits by province (Million TL)	X_1: Savings deposit X_2: Official institutions deposits X_3: Commercial organizations deposits X_4: Foreign currency deposit account X_5: Other institutions deposit X_6: Precious metals deposit accounts	Y_1: Savings deposits per capita (TL) Y_2: Official institutions deposits per capita (TL) Y_3: Deposits of commercial organizations per capita (TL) Y_4: Foreign exchange deposit account per capita (TL) Y_5: Other institutions deposits per capita (TL) Y_6: Precious metals deposit accounts per capita (TL)
Loan by province (Million TL)	X_7: Ağrı (04)cultural specialization loan X_8: Real estate specialty loan X_9 Vocational specialization loan X_{10}: Tourism specialization loan X_{11}: Other specialty loans X_{12} :Non-specialized loans	Y_7: Ağrı (04)cultural specialization loan per capita (TL) Y_8: Real estate specialty loans per capita (TL) Y_9: Vocational specialization credit per person (TL) Y_{10}: Tourism specialization credit per capita (TL) Y_{11}: Other specialized loans per capita (TL) Y_{12}: Non-specialized loans per capita (TL)
Population	N: Population in 2022	

The data show the distribution of deposits and loans by province as of 31 December 2022. Since the distribution of deposits and loans in densely populated provinces will not be homogeneous compared to other provinces, it was considered useful to weight the existing data by population. The 2022 population data by province was obtained from the Turkish Statistical Institute website (URL-3). With this weighting, the new variables were reorganised as bank deposits per capita and bank loans per capita. The new variables are also shown in Table 5.

In this study, per capita bank deposits ($Y_1 - Y_6$) and per capita bank loans ($Y_7 - Y_{12}$) were calculated based on 2002 population data by province, and the clustering of provinces was evaluated using these data. The number of clusters was determined as six, based on the level of development ranking specified in the SEGE 2017 working report prepared by the General Directorate of Development Agencies of the Ministry of Industry and Technology of the Republic of Turkey (SEGE, 2017). Taking the number of clusters as six is also supported by the $k \cong (n/2)^{1/2}$ relationship, which is known as a practical method for determining the number of clusters (Tatlıdil, 2002). Since $n = 81$, the approximate number of clusters is determined as $k = \sqrt{(81/2)} = 6{,}36 \cong 6$.

Results

In this study, provinces in Turkey were classified using different clustering methods based on the variables of bank deposits per capita and bank loans per capita. The clustering structures of the provinces were determined using the complete linkage clustering method, the median clustering method, the Ward clustering method, and the K-means clustering method. This allows for an evaluation of the behaviour of clustering methods on the same data.

Separate classifications were made for the provinces' bank deposits and bank loans. The reason for this is to mutually evaluate the clustering status of the provinces in terms of deposits and loans. When applying clustering methods, the number of clusters was set to k=6 for reasons stated earlier.

Some Findings Based on The 2017 SEGE Working Report

Based on the findings obtained in the 2017 Provincial SEGE working report, the distribution of provinces according to the socio-economic development scale was organised in Table 4. Some statistics on various bank deposits and various bank loans per capita for provinces according to the 2017 SEGE study are provided in Table 6.

Table 6 shows that provinces in the first tier, according to the classification established based on the 2017 SEGE study report, have higher averages than provinces in other tiers in terms of both deposits and loans per capita. It is possible to say that this result stems from the fact that the main tiers of the central administration, advanced industrialisation, and provinces ranked first in terms of socio-economic development are included in this tier.

The sixth tier generally includes provinces located in the Eastern and Southeastern Anatolia regions, which are weak in terms of industrialisation and have low socio-economic development rankings. Therefore, it is observed that provinces in the sixth tier also have the lowest average per capita bank deposits and bank loans. The number of tiers established according to the 2017 Provincial SEGE working report was taken as the appropriate number of clusters, supported by an examination of whether there was a significant difference between the groups (tiers) in terms of per capita bank deposits and bank loans. For this purpose, multivariate analysis of variance (MANOVA), which is used to compare the mean vectors of more than two independent groups, was employed.

Tablo 6. Some statistics for grades created according to the 2017 Provincial SEGE

Stage	Variable		Mean	S.D.	Stage	Variable		Mean	S.D.
1	Deposit	Y_1	35015,84	10284,48	2	Deposit	Y_1	21486,40	6406,67
		Y_2	5698,73	13670,08			Y_2	808,38	315,97
		Y_3	20352,99	17284,71			Y_3	6485,76	3072,50
		Y_4	47312,13	26310,04			Y_4	16479,73	6343,55
		Y_5	2335,08	3525,75			Y_5	892,99	400,74
		Y_6	6373,26	1533,25			Y_6	4995,42	941,58
	Loan	Y_7	2964,88	1526,88		Loan	Y_7	4711,93	2076,74
		Y_8	315,84	137,92			Y_8	351,44	145,15
		Y_9	11,07	3,55			Y_9	10,71	5,74
		Y_{10}	449,39	449,33			Y_{10}	527,58	726,37
		Y_{11}	5837,08	13779,75			Y_{11}	10788,29	16243,55
		Y_{12}	90938,73	44793,20			Y_{12}	43728,99	18644,39
3	Deposit	Y_1	17753,89	5411,18	4	Deposit	Y_1	16384,66	3011,82
		Y_2	845,38	388,23			Y_2	1033,48	773,16
		Y_3	6238,35	2867,23			Y_3	4416,83	2709,30
		Y_4	16258,33	6169,15			Y_4	16090,49	7789,26
		Y_5	751,33	391,23			Y_5	494,78	215,84
		Y_6	4398,09	1199,85			Y_6	3944,28	823,60
	Loan	Y_7	3606,80	2260,69		Loan	Y_7	5644,32	3511,94
		Y_8	266,51	67,75			Y_8	358,45	171,01
		Y_9	12,26	6,06			Y_9	13,35	8,35
		Y_{10}	106,90	5920			Y_{10}	113,52	91,55
		Y_{11}	1512,28	469,88			Y_{11}	1699,66	761,98
		Y_{12}	39673,71	17649,73			Y_{12}	27564,58	6676,46
5	Deposit	Y_1	12723,09	5720,44	6	Deposit	Y_1	6228,31	2561,78
		Y_2	1021,56	544,82			Y_2	1106,66	430,78
		Y_3	2757,09	737,07			Y_3	1935,29	498,14
		Y_4	10643,44	7121,20			Y_4	3469,11	2633,91
		Y_5	420,54	227,14			Y_5	208,96	65,21
		Y_6	3177,82	852,10			Y_6	1449,11	534,44
	Loan	Y_7	4722,76	2593,24		Loan	Y_7	3169,66	2451,55
		Y_8	320,33	122,03			Y_8	153,10	91,85
		Y_9	15,77	7,77			Y_9	9,21	6,79
		Y_{10}	105,26	55,63			Y_{10}	47,42	25,20
		Y_{11}	1802,05	693,72			Y_{11}	755,55	495,68
		Y_{12}	24797,27	7398,71			Y_{12}	15218,57	4701,43

If there is a statistically significant difference between the group mean vectors in terms of all variables, then the number of independent groups can be taken as the appropriate cluster number, as the groups will be well separated from each other. MANOVA results for both bank deposits and bank loans are given in Table 7.

According to Table 7, there is a statistically significant difference between the groups (levels) in terms of both bank deposits per capita and bank loans per capita ($p < 0.001$). This result supports our decision to take the number of clusters as $k = 6$, as determined in Section 2.2, clustering of provinces based on both bank deposits per capita and bank loans per capita. Thus, the maximum number of clusters applied in the clustering of provinces according to the aforementioned criteria was set at 6. Accordingly, the cases where the number of clusters was $k = 4,5$ and 6 were evaluated using clustering methods.

Tablo 7. Bank deposits and loans per capita MANOVA result

Factor		Method	Statistic	F	(sd_1, sd_2)	p
Stage	Deposit	Pillai's Trace Test	1,304	4,350	(30 ,370)	0,000***
		Wilks' Lambda Test	0,117	6,688	(30 ,282)	0,000***
		Hotelling's Trace Test	4,322	9,854	(30 ,342)	0,000***
		Roy's Greatest Eigenvalue Test	3,520	43,414	(6 ,74)	0,000***
	Loan	Pillai's Trace Test	1,096	3,461	(30 ,370)	0,000***
		Wilks' Lambda Test	0,217	4,377	(30 ,282)	0,000***
		Hotelling's Trace Test	2,336	5,326	(30 ,342)	0,000***
		Roy's Greatest Eigenvalue Test	1,743	21,496	(6 ,74)	0,000***

Complete Linkage Clustering Method Result and Evaluation

In this section, provinces were classified according to per capita bank deposits and per capita bank loans using complete linkage clustering method with cluster numbers k = 4,5 and 6. The results are presented in Appendix (App.) Table 1 and evaluated to determine the appropriate clustering structure.

According to App. Table 1, regardless of the number of clusters, Istanbul and Ankara (06) are alone in the first and second clusters in terms of bank deposits per capita. This can be explained by the fact that institutions affiliated with the central administration and financial centres are predominantly concentrated in these two provinces. When the number of clusters is taken as 4, 5, and 6, it is generally observed that the provinces in the first three clusters do not change.

While the first and second clusters consist of a single province, the third cluster covers four provinces. In other clusters, provinces are observed to change places between clusters. Another noteworthy point is that when the number of clusters is 5 or 6, the first three clusters and the last clusters consist of the same provinces. However, the provinces in the fourth cluster obtained when $k = 5$ are divided into two within themselves to form new clusters when $k = 6$. In terms of bank deposits per capita, it is noteworthy that the provinces of Aksaray (68) and Tunceli (62) are in the cluster with the provinces of Kocaeli (41), Denizli (20) and Kayseri (38) when $k = 5$ and $k = 6$, that most provinces are grouped in the last clusters, and that provinces such as Konya (42), Gaziantep (27) and Samsun (55) are in the last clusters. A graphical representation of the results of the complete linkage clustering method, called a dendrogram, is given in Figure 1.

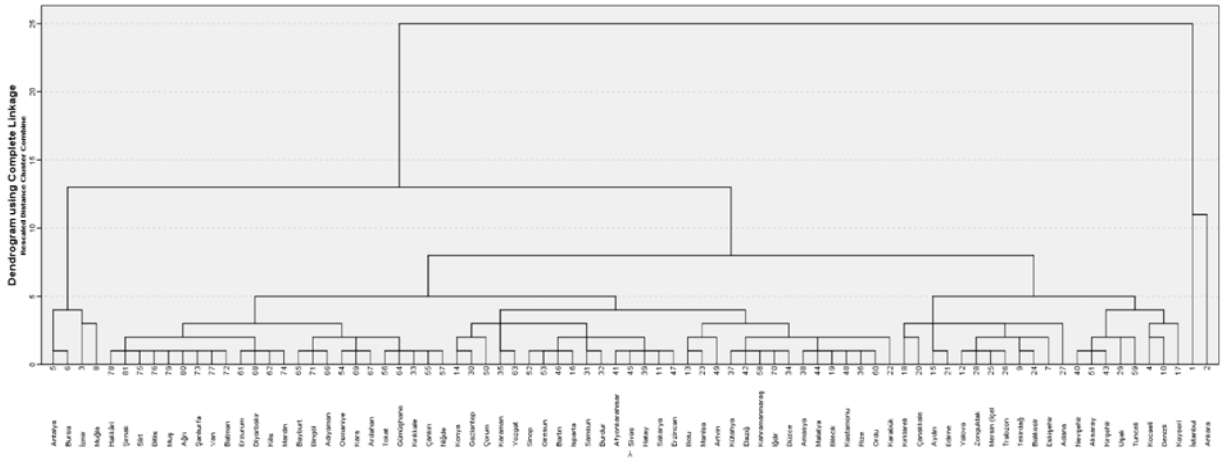

Figure 1. Dendrogram of the complete linkage clustering method according to per capita bank deposits in provinces in Turkey

Considering that the final clusters consist of the same provinces when the number of clusters is set to 5 or 6, and considering Figure 1, it is possible to say that the best clustering of provinces in terms of bank deposits per capita in the complete linkage method is when $k = 5$.

The results of classifying Turkish provinces according to the number of clusters obtained using the complete linkage clustering method based on per capita bank loans are presented in App. Table 2, while the dendrogram is shown in Figure 2. According to App. Table 2, even if the number of clusters changes, the first and second clusters remain unchanged in terms of bank loans per capita, with the same provinces included in these clusters. The vast majority of provinces are in the fourth cluster when $k = 4$, in the fifth cluster when $k = 5$, and again in the fifth cluster when $k = 6$. When the number of clusters is 4 or 5, they consist of the same provinces. When the number of clusters is 4, 5 provinces are separated from the third cluster to form the fourth cluster when $k = 5$, while the remaining 60 provinces form the fifth cluster. Furthermore, while the first four clusters for $k = 5$ and $k = 6$ consist of the same provinces, the fifth cluster obtained for $k = 5$ splits into two, forming the fifth and sixth clusters for $k = 6$. Regardless of the number of clusters, it is meaningful that İstanbul (34) and Ankara (06) always rank in the first cluster in terms of bank loans per capita.

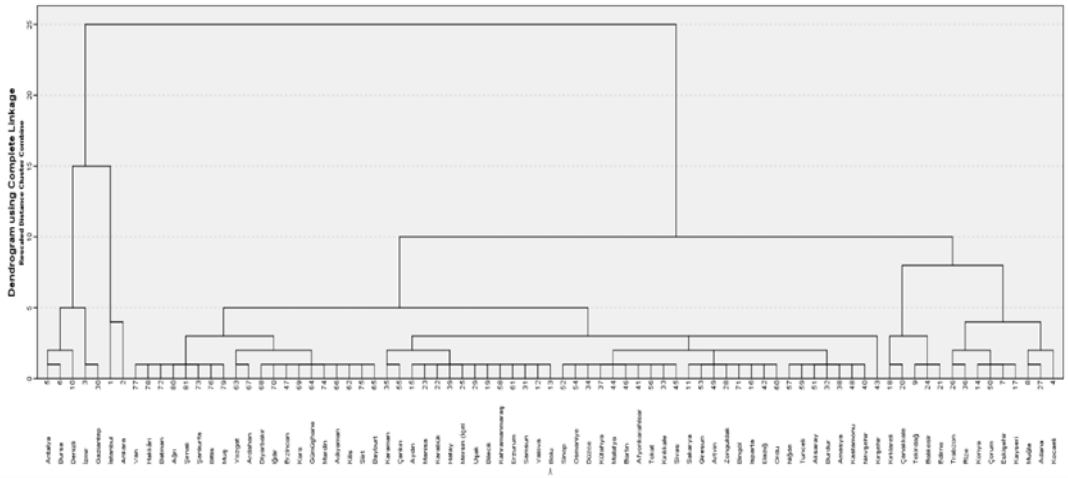

Figure 2: Dendrogram of the complete linkage clustering method according to per capita bank loans in provinces of Turkey

On the other hand, when the number of clusters is taken as 5 and 6, the fact that the provinces in the first three clusters generally remain unchanged can be interpreted as the provinces in these clusters being similar in terms of bank loans per capita. It is noteworthy that Rize (53) and Çorum (19) are in the same cluster as provinces with developed industries or important tourism destinations, such as Kocaeli (41), Kayseri (38), and Muğla (48), in terms of bank loans per capita. When the clustering structures based on the cluster numbers given in App. Table 2 and the dendrogram given in Figure 2 are evaluated together, it can be said that classifying the provinces of Turkey into $k = 5$ clusters in terms of bank loans per capita is more appropriate.

Median Clustering Method Results and Evaluation

The classification of provinces in Turkey according to bank deposits per capita was examined using the median clustering method with cluster numbers $k = 4,5$ and 6. The classification results according to bank deposits were evaluated and presented in App. Table 3, and appropriate clustering structures were determined.

According to App. Table 3, regardless of the number of clusters, İstanbul (34) and Ankara (06) are alone in the first and second clusters, respectively, in terms of bank deposits per capita. When the number of clusters is taken as 4, only İzmir (35), Antalya (07), Bursa (16), and Muğla (48) form the third cluster, separate from the other provinces, while the other provinces form the fourth cluster. There is a bottleneck in the last group. When examining the clustering structures formed when the number of clusters is set to 5 or 6, it is seen that in both cases, the last clusters consist of the same provinces. The only difference is that when $k = 6$, Kocaeli (41) and Denizli (20), which are in the fourth cluster when the number of clusters is 5, are separated and assigned to the fourth cluster, while the remaining provinces are assigned to the fifth cluster. Therefore, it would be more appropriate to use five clusters when classifying Turkish provinces based on per capita bank deposits using the median method. This is also evident in the dendrogram shown in Figure 3.

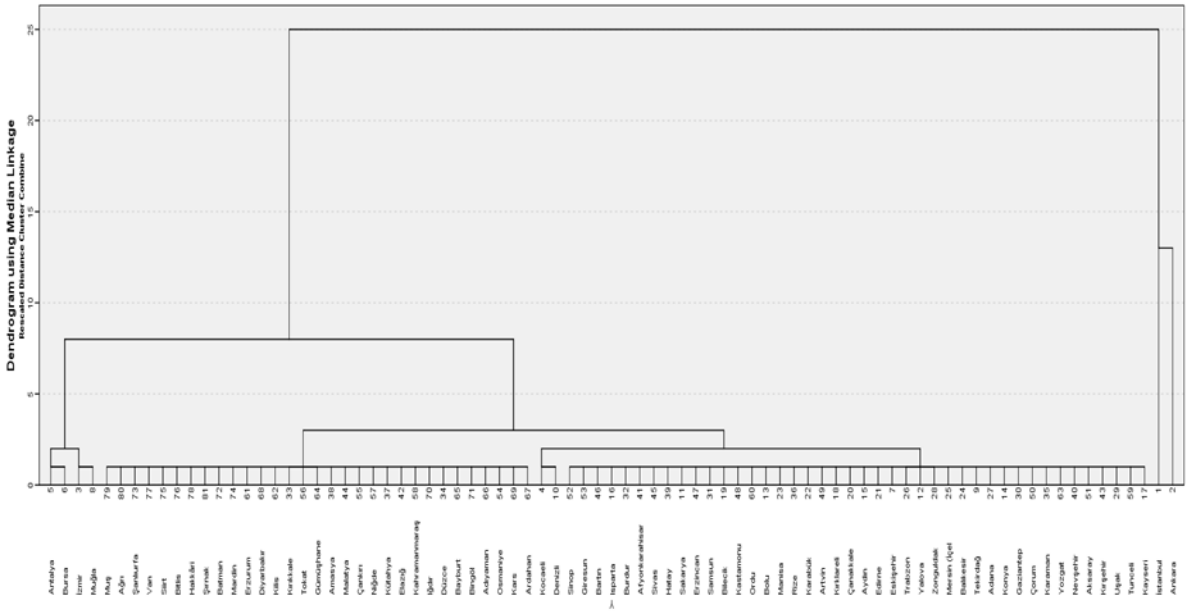

Figure 3. Dendrogram of the median clustering method for provinces in Turkey based on bank deposits per capita

The results of the classification of provinces in Turkey based on bank loans per capita using the median clustering method for $k = 4, 5,$ and 6 are shown in App. Table 4. According to these results, it would be appropriate to take the number of clusters as $k = 4$ or $k = 5$. This is because, although there is an accumulation in the last clusters in each case, in the $k = 4, 5,$ and 6 cases, the last clusters always consist of the same provinces. Furthermore, in the case of $k = 5$, the second and third clusters are formed by splitting the second cluster in the case of $k = 4$. Similarly, in the case of $k = 6$, the first and second clusters are formed by splitting the first cluster in the case of $k = 5$ and comprise only the provinces of İstanbul (34) and Ankara (06), respectively. The classification of provinces in terms of bank loans per capita using the complete linkage and median methods for $k = 5$ yields the same clusters. Therefore, it would be more accurate to use five clusters when classifying Turkish provinces according to bank loans per capita using the median method. This is supported by the dendrogram shown in Figure 4.

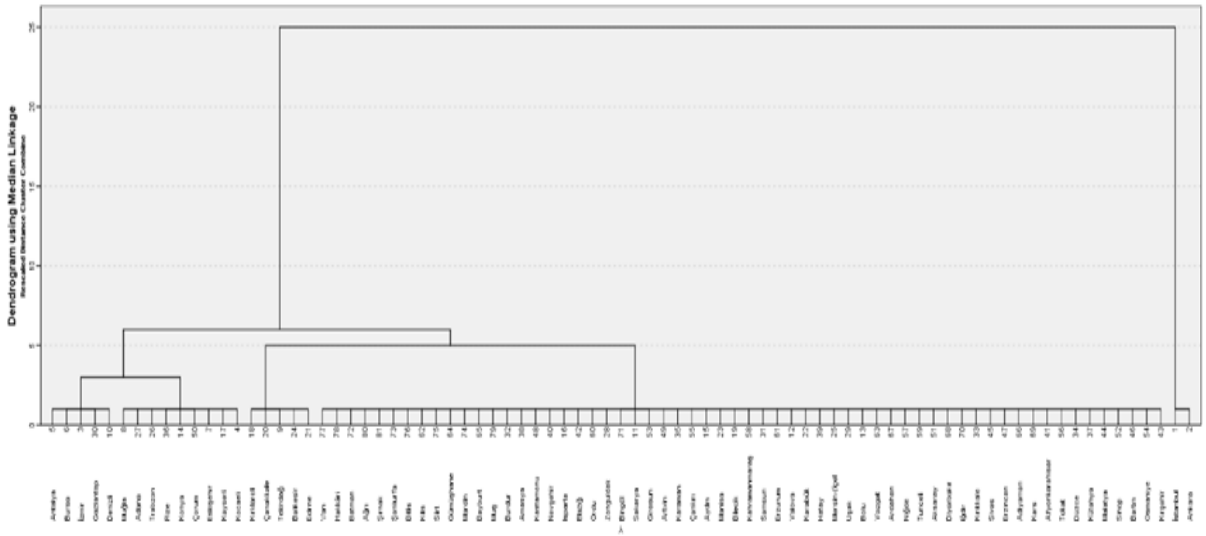

Figure 4. Median clustering method dendrogram of provinces in Turkey according to bank loans per capita

Results and evaluation of the ward clustering method

This section evaluates the classification of provinces in Turkey according 475op er capita bank deposits using the Ward clustering method. The results of the classification study, with cluster numbers $k = 4,5$ and 6, are presented in App. Table 5.

As with other clustering methods, in the Ward method, İstanbul (34) and Ankara (06) form either a single cluster or separate clusters depending on the varying number of clusters in the clustering of provinces according 475op er capita bank deposits. As the number of clusters increases in the Ward method, the resulting clustering structure resembles that of other methods. When the number of clusters is $k = 4$ and 5, the first three clusters consist of the same provinces. When the number of clusters is 4, the fourth cluster is formed, and when the number of clusters is 5, it is divided into two, resulting in the fourth and fifth clusters. When the number of clusters is 5 and 6, the same provinces are assigned to the last three clusters. The only difference between the two situations is that in the $k = 5$ situation, the provinces of İstanbul (34) and Ankara (06), which are in the first cluster, are separated in the $k = 6$ situation and form separate clusters.

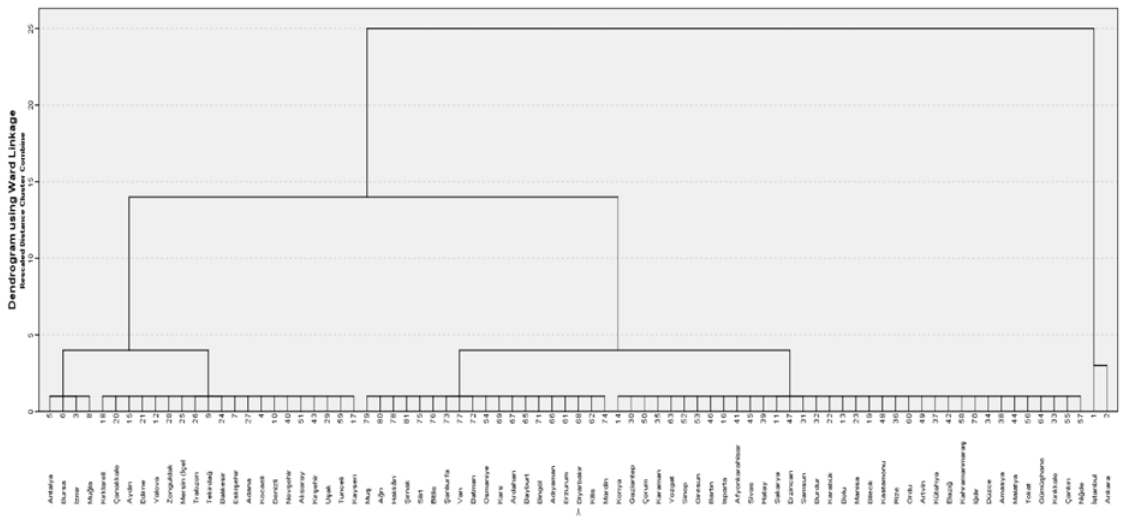

Figure 5. Dendrogram of Turkish provinces according to per capita bank deposits using the Ward clustering method

Consequently, it would be appropriate to use five clusters when classifying Turkish provinces according to per capita bank deposits using the Ward clustering method. This is also evident in the dendrogram shown in Figure 5.

The results of classifying Turkish provinces according to bank loans per capita using the Ward clustering method are organised in App. Table 6 according to varying numbers of clusters. Regardless of the number of clusters, the first cluster consisting of İstanbul (34) and Ankara (06) provinces remains unchanged. The validity of this in the Ward method can also be seen by looking at the dendrogram shown in Figure 6.

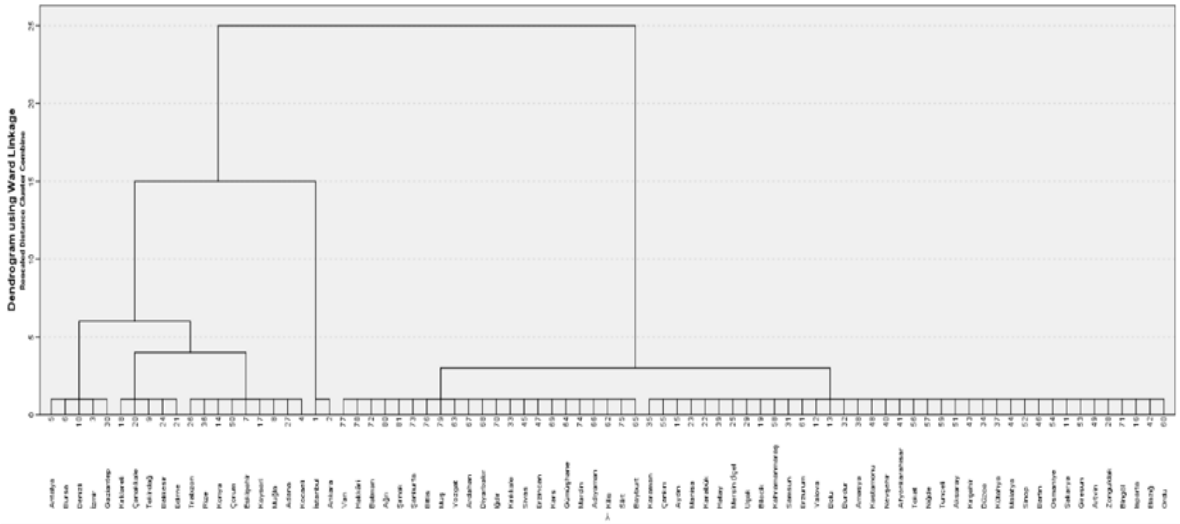

Figure 6. Ward clustering method dendrogram according to bank loans per capita of provinces in Turkey

K-Means clustering method results and evaluation

This section evaluates the classification of Turkish provinces based on per capita bank deposits and bank loans using the K-means clustering method. The results of the classification study based on per capita bank deposits, with cluster numbers $k = 4, 5$ and 6, are presented in App. Table 7. Regardless of the number of clusters in the clustering, İstanbul (34) and Ankara (06) alone form the first and second clusters, respectively. Furthermore, when the number of clusters is four, the provinces of Kocaeli (41) and Denizli (20), which are classified into the third cluster, are classified into the fourth cluster when the number of clusters is 5 and 6. In both cases, the provinces classified into the third cluster remain the same. When the number of clusters is 4, the vast majority of provinces are clustered in the fourth cluster, which is generally not a statistically preferred situation. However, the provinces classified into the fourth and fifth clusters when $k = 5$ and those classified into the fifth and sixth clusters when $k = 6$ show a high degree of similarity. For this reason, it is appropriate to use a cluster number of 5 when classifying Turkish provinces according to per capita bank deposits using the K-means method.

The results of the classification of provinces in Turkey according to per capita bank loans using the K-means clustering method are presented in App. Table 8, organised according to the varying number of clusters. Number of clusters Regardless of the number of clusters in the clustering, İstanbul (34) and Ankara (06) together form the first cluster.

Furthermore, the provinces classified into the second cluster when the number of clusters is 4, the provinces classified into the third cluster when the number of clusters is 5, and the provinces classified into the sixth cluster when the number of clusters is 6 are the same provinces. When the number of clusters is 4, the vast majority of provinces are clustered in the

fourth cluster, which is an undesirable situation in cluster analyses. When the number of clusters is set to 5 and 6, the provinces classified into the fourth cluster remain the same. However, the provinces classified into the second and fifth clusters when $k = 5$ and the provinces classified into the second and sixth clusters when $k = 6$ show a high degree of similarity. For this reason, it would be more appropriate to again take the number of clusters as 5 when classifying the provinces of Turkey according to bank loans per capita using the K-means method

The cluster centres or averages of variables according to groups (clusters) obtained as a result of classification analyses performed using four separate clustering methods based on bank deposits per capita and bank loans per capita in Turkish provinces are presented in Table 8 and Table 9, respectively, in relation to the relevant topics.

The clusters in the rows shown in the same colour in Tables 8 and 9 contain the same provinces. Therefore, the cluster centres in the rows of the same colour are equal to each other according to the methods. Since the clusters shown in yellow contain a single province (İstanbul (34) in the first cluster, Ankara (06) in the second cluster), the variable values for these provinces are also their own cluster centres. In the clusters in the colourless rows, the provinces are generally similar, but depending on the clustering methods, there are shifts in the provinces between clusters. For this reason, the cluster centres of these clusters differ according to the methods. The clusters in the rows shown in the same colour in Tables 8 and 9 contain the same provinces. Therefore, the cluster centres in the rows of the same colour are equal to each other according to the methods. Since the clusters shown in yellow contain a single province (İstanbul (34) in the first cluster, Ankara (06) in the second cluster), the variable values for these provinces are also their own cluster centres. In the clusters in the colourless rows, the provinces in the clusters are generally similar, but depending on the clustering methods, there are shifts in the provinces between clusters. For this reason, the cluster centres of these clusters differ according to the methods.

Table 8. Cluster centres of clusters obtained based on per capita bank deposits

Method	Cluster No	Number of Province	Cluster centers (Deposit averages per capita by groups)					
			Y_1	Y_2	Y_3	Y_4	Y_5	Y_6
Complete linkage method	1	1	53077,80	1898,23	57939,89	97444,23	3162,00	8867,45
	2	1	34236,64	42140,95	40248,62	77384,63	11488,54	8725,10
	3	4	38758,09	980,74	12018,97	46104,21	744,21	6043,55
	4	20	23954,65	864,38	7151,73	22591,30	884,46	4601,51
	5	55	11890,19	1011,48	3702,95	9146,77	454,89	3246,62
Median	1	1	53077,80	1898,23	57939,89	97444,23	3162,00	8867,45
	2	1	34236,64	42140,95	40248,62	77384,63	11488,54	8725,10
	3	4	38758,09	980,74	12018,97	46104,21	744,21	6043,55
	4	44	20013,33	912,83	6187,70	17723,40	749,37	4460,91
	5	31	8144,11	1056,60	2401,24	5647,38	314,06	2397,23
Ward	1	2	43657,22	22019,59	49094,26	87414,43	7325,27	8796,28
	2	4	38758,09	980,74	12018,97	46104,21	744,21	6043,55
	3	20	23954,65	864,38	7151,73	22591,30	884,46	4601,51
	4	36	14880,73	927,96	4580,01	12170,03	564,16	4109,18
	5	19	6223,93	1169,73	2041,17	3418,49	247,85	1612,29
K-means	1	1	53077,80	1898,23	57939,89	97444,23	3162,00	8867,45
	2	1	34236,64	42140,95	40248,62	77384,63	11488,54	8725,10
	3	27	22570,13	868,36	6298,73	21235,00	768,69	4640,24
	4	48	10909,59	1030,70	3679,82	7949,03	457,37	3027,25
	5	4	38758,09	980,74	12018,97	46104,21	744,21	6043,55

Table 9. Cluster centres of clusters obtained based on per capita bank loans

Method	Cluster Number	Number of Province	Cluster centers (Credit averages per capita by groups)					
			Y_7	Y_8	Y_9	Y_{10}	Y_{11}	Y_{12}
Complete linkage Median Ward	1	2	1295,78	226,76	8,57	185,72	826,08	163417,85
	2	5	3593,57	396,83	13,15	433,88	1672,96	91782,42
	3	9	3575,85	237,07	8,89	108,50	1266,13	52081,59
	4	5	6086,30	429,48	11,18	1574,79	37764,69	46064,45
	5	60	4282,19	280,21	12,60	96,30	1414,91	24723,59
K-means	1	2	1295,78	226,76	8,57	185,72	826,08	163417,85
	2	15	3782,40	297,66	10,63	109,73	1368,97	46719,18
	3	5	6086,30	429,48	11,18	1574,79	37764,69	46064,45
	4	54	4303,30	268,17	12,53	94,61	1402,88	23173,37
	5	5	3593,57	396,83	13,15	433,88	1672,96	91782,42

Conclusion and Discussion

The Statistical Regional Units Classification (SRUC), established in 2001 by the Turkish Statistical Institute (TUIK) and the State Planning Organisation (DPT), entered into force with the publication of the Council of Ministers' Decision No. 2002/4720 in the Legal Gazette dated 22 September 2022. The objectives of preparing the SRUC, which Turkey, seeking to join the European Union (EU), prepared taking into account EU requirements, can be listed as follows: To collect and develop regional statistics, conduct socio-economic analyses of regions, determine the framework for regional policies, and establish basic indicators that will contribute to the creation of a comparable statistical database compatible with the EU Regional Statistical System (Erilli, 2014b).

Within the framework of the objectives of preparing the IBBS, classification or clustering studies of provinces based on various criteria play an important role. Perhaps the most important of these studies, or the one that could guide other studies, is the 2017 Provincial SEGE study prepared by the General Directorate of Development Agencies of the Ministry of Industry and Technology of the Republic of Turkey. As a result of the classification of provinces according to their socio-economic development ranking in the 2017 Provincial SEGE, provinces were divided into 6 levels (Table 4).

In this study, using TBB's 2022 bank deposit (6 different deposits) and bank loan (6 different loans) data and TUIK's 2022 population data, bank deposits per capita and bank loans per capita were first calculated for 81 provinces. Then, using the calculated data, the 81 provinces were clustered together based on their similarity in terms of both bank deposits and bank loans per capita. The purpose of classifying the provinces separately according to bank deposits and bank loans was to see how the similarity structure of the provinces changed according to deposits and loans. Both hierarchical and non-hierarchical clustering methods were used in the clustering of provinces. Among the hierarchical methods, the complete linkage, median, and Ward clustering methods were used, while the K-means clustering method was used among the non-hierarchical methods.

In the study, the number of clusters was accepted as the maximum possible number of clusters based on the number of levels created according to the socio-economic development scale of the provinces in the 2017 Provincial SEGE, and the behaviour of the clustering methods was examined when the number of clusters was 4, 5, and 6. The evaluations concluded that the most appropriate clustering structure for all clustering methods was achieved when the number of clusters was 5. When the number of clusters was 4, it was observed that the vast majority of

provinces were grouped into a single cluster in both deposit-based and loan-based clustering. When the number of clusters was 6, the analysis of bank deposits in particular showed that İstanbul (34) and Ankara (06) formed separate clusters in all clustering methods. This situation only emerged in the median method in the classification analysis of bank loans. Furthermore, when the number of clusters is 6, the provinces classified into the sixth cluster are mostly the same as those classified into the fifth cluster when the number of clusters is 5. For these reasons, the number of clusters adopted in all clustering methods is 5.

When the number of clusters is five, it can be said that the best classification for provinces in terms of bank deposits per capita is obtained using the Ward method. This is because in other methods, there is one province (İstanbul (34) and Ankara (06)) in each of the first two clusters. In the Ward method, however, these two provinces are classified into the first cluster. The reason for this can be attributed to the concentration of administrative mechanisms and institutions, as well as financial institutions, in these two provinces, which are responsible for making decisions on economic, social, and administrative matters. This situation can be considered an indicator of the concentration of bank deposits in these provinces. The second cluster in the Ward method and the third cluster in other hierarchical methods, as well as the fifth cluster in the K-means method, consist of the same provinces. As these provinces are primarily prominent for tourism, it can be said that the concentration of bank deposits in these areas is of secondary importance. The fifth clusters mainly consist of provinces in the Southeast and East Anatolia regions, although it is possible to see that some provinces from other regions are also classified in this cluster.

When the number of clusters is 5, it can be said that the best classification for provinces in terms of bank loans per capita is obtained using the K-means method. In reality, all methods yield almost the same classification. This is because the first, second, and fourth clusters in hierarchical methods correspond to the first, fifth, and third clusters in the K-means method, consisting of the same provinces. The only difference is in the other clusters.

When the number of clusters is 5, it is possible to state that the best classification for provinces in terms of bank loans per capita is obtained using the K-means method. In reality, all methods yield almost the same classification. This is because the first, second, and fourth clusters in hierarchical methods correspond to the first, fifth, and third clusters in the K-means method, consisting of the same provinces. Differences are only observed in the other clusters. In all methods, İstanbul (34) and Ankara (06) are classified into one cluster. Similarly, İzmir (35), Antalya (07), Bursa (16), Denizli (20) and Gaziantep (27) are classified into one cluster, while Tekirdağ (59), Kırklareli (39), Çanakkale (17), Edirne (22) and Balıkesir (10) provinces are classified into a separate cluster. In this classification based on bank lending, the primary importance of İstanbul (34) and Ankara (06) can be attributed to decision-making mechanisms and financial resources. Tourism and industrial development can be considered as the sources of the second-degree importance of İzmir (35), Antalya (07), Bursa (16), Denizli (20), and Gaziantep (27) in the classification analysis based on bank lending. In the last clusters, where provinces are densely located, it is seen that provinces from other regions, especially those in the Southeast and Eastern Anatolia regions, are clustered in all methods.

When the findings of this study are compared with the results of other studies, there are both similarities and differences. Even when the focus of the studies is the same, differences in the variables used affect the units to be classified into clusters. When identifying the clustering structures of provinces in Turkey, it is not expected that the results of studies using socio-economic variables will be similar to those using variables related to bank deposits or bank loans. This situation can be observed in our study as well as in the works of Karakaş and Erilli (2023) and Karabulut et al. (2004).

Yılmaz and Uzgören (2013) classified Turkey's provinces into six clusters using the K-means method based on banking activities. Although they used different variables than those we employed, the results obtained generally resemble our findings when the number of clusters in our K-means analysis was set to six. This similarity is particularly striking in the classification results based on per capita bank deposits. The differences observed in some provinces within certain clusters stem from the variations in the variables used in the studies.

Sarıgül (2014) examined the clustering of provinces in terms of access to and use of banking and financial services. In this study, where the number of clusters was set to 6 and a hierarchical clustering method was used in both cases, the distribution of provinces across clusters differs from the findings.

In their study on the clustering of provinces in Turkey using the K-means method based on credit usage levels, Tekin and Temelli (2020) clustered the provinces into five-year periods between 1988 and 2018. This study differs from ours both in terms of the variables used and its objectives. For this reason, the cluster structures emerging from the results of the K-means method used in our study are quite different from the cluster structures obtained for all periods in that study. While the number of clusters in that study could be 9, 10, or 12, in our study it is 5 or 6.

Yaman Yılmaz (2020) conducted a classification of provinces in Turkey in terms of banking activities using different clustering methods based on 10 variables. While this study resembles our own in terms of the clustering methods used, it differs in terms of the variables employed. In this study, separate clustering analyses were performed using both the raw data and the principal component data obtained by applying principal component analysis to the raw data, and the results were evaluated. In this respect, the study in question also differs from our own. In the clustering resu lts obtained using both raw data and principal component data, when the number of clusters was set to 2 or 3, it was observed that nearly all provinces were grouped into a single cluster. In our study, however, the provinces were divided into 5 or 6 clusters, and a balanced distribution was observed for provinces other than the major ones, depending on the clustering methods.

In conclusion, according to the findings obtained from this study, the clustering structures of provinces in terms of bank deposits and bank loans per capita are not consistent with the tier structure established according to the socio-development scale of provinces in the 2017 Provincial SEGE. Whether the differences between groups in the clustering structure observed in terms of bank deposits are significant can be investigated using the Ward method, while in the Complete Linkage, Median, and K-Means methods, it can be done on the other three clusters after the first and second clusters are extracted. In terms of bank loans, the significance of the difference between groups can be investigated for all methods. As revealed by this study, there are significant differences between provinces in terms of the clusters formed for various types of bank deposits and bank loans in our country. More detailed analyses can determine on which variables these differences are or will be more pronounced. Furthermore, as a suggestion, the appropriateness of the number of clusters can be evaluated using cluster validity indices. This study is expected to guide those operating in the banking sector in particular towards undertaking work aimed at eliminating the differences between provinces in terms of bank deposits and loans per capita.

Author contribution ratio

All processes of this study (design, data collection, analysis, writing, and editing) were carried out by Burçin Öner.

Ethics committee approval

Ethics committee approval is not required for the prepared study.

Declaration of conflict of interest

There is no conflict of interest with any person/institution in the prepared study.

Article Process History

Article Submission Date: 10.02.2026
Article Acceptance Date: 06.04.2026

References

Açıkalın; S. (2023). *Covid-19 Pandemi Döneminde Türk Bankacılık Sektöründe Yer Alan Bankaların Kümeleme Analizi İle Sınıflandırılması,* (Yayımlamamış Yüksek Lisans Tezi), Balıkesir Üniversitesi, Balıkesir.

Aggarwal, C.C. & Reddy, C.K. (2014). *Data Clustering Algorithms and Applications*. CRC Press Taylor&Francis Group, New York.

Akgöz, E. (2010). *Türkiye'de Ticari Bankaların Finansal Oranlar Yardımıyla Sınıflandırılması: Kümeleme Analizi Yaklaşımı*, (Yayımlanmamış Yüksek Lisans Tezi), Kahramanmaraş Sütçü İmam Üniversitesi, Kahramanmaraş.

Albayrak, A. S. & Savaş, F. (2015). Türkiye'de İllerin Sosyo-Ekonomik Gelişmişliğinin Belirleyicileri ve 2012 Yılı Sosyo-Ekonomik Gelişmişlik Sıralaması. *AİBÜ Sosyal Bilimler Enstitüsü Dergisi*, 15(3), 1-40.

Alpar, R. (2020). *Uygulamalı Çok Değişkenli İstatistiksel Yöntemler* (6th Ed.). Detay Yayıncılık, Ankara.

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

Aydın, D. & Başkır, M. B. (2013). Bankaların 2012 Yılı Sermaye Yeterlilik Rasyolarına Göre Kümeleme Analizi ve Çok Boyutlu Ölçekleme Sonucu Sınıflandırılma Yapıları. *Bankacılık ve Sigortacılık Araştırmaları Dergisi*, 1(5-6), 29-47. https://doi.org/10.1501/bsad_0000000015

Ayrıçay, Y. & Akgöz, E. (2014). Ticari Bankaların Finansal Oranlar Yardımıyla Sınıflandırılması: Kümeleme Analizi Yaklaşımı. *Journal Of Social and Humanities Sciences Research*, 1(1), 1-23. https://doi.org/10.26450/jshsr.5

Bulut, H. (2018). *R Uygulamaları ile Çok Değişkenli İstatistiksel Yöntemler*. Nobel Akademik Yayıncılık.

Bulut, H. (2019). Türkiye'deki İllerin Yaşam Endekslerine Göre Kümelenmesi. *Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 23(1), 74-82, 2019. https://doi.org/10.19113/sdufenbed.444143

Çağırgan, T. S. & Ulusoy, Y. (2019). Kredi ve kredi işlemlerinin hukuki açıdan değerlendirilmesi. *PJESS*, 6(2), 43-69, 2019. https://doi.org/10.34232/pjess.572449

Er. B. (2022). Mevduat Sözleşmesinin Tanımı, Kurulması ve Türleri. *Türkiye Adalet Akademi Dergisi*, 13(49), 491-516. https://doi.org/10.54049/taad.1055809

Erilli, N. A. (2014a). TR72 Bölgesi İlçelerinin Sosyo-Ekonomik Verilere Göre Bulanık Kümeleme Analizi ile Sınıflandırılması. *The International Journal of Economic and Social Research*, 10(2), 33-45.

Erilli, N. A. (2014b). Bulanık Kümeleme Analizi ile İstatistiki Bölge Birimlerinin (İBBS) Mali Değişkenlere Göre Sınıflandırılması. *Kırıkkale Üniversitesi Sosyal Bilimler Dergisi*, 4(2), 149-165.

Kalaycı, Ş. (2005). *SPSS Uygulamalı Çok Değişkenli İstatistik Teknikleri*. Asil Yayıncılık, Ankara.

Karaatlı, M. & Yıldız, E. (2021). Mevduat Bankaların Finansal Yapılarının Kümeleme Analizi ile İncelenmesi. *BMIJ*, 9(1), 1-17. https://doi.org/10.15295/bmij.v9i1.1594.

Karabulut, M., Gürbüz, M. & Sandal, E. K. (2004). Hiyerarşik Kluster (Küme) Tekniği Kullanılarak Türkiye'de İllerin Sosyo-Ekonomik Benzerliklerinin Analizi. *Coğrafi Bilimler Dergisi*, 2(2), 65- 78. https://doi.org/10.1501/Cogbil_0000000043

Karadaş, K. & Erilli, N. A. (2023). Gri Kümeleme Analizi ile Türkiye'deki Şehirlerin Sosyo-Ekonomik Verilere Göre Sınıflandırılması. *Journal of Statistical Research*, 13(2), 60-74.

Karaduman, M. (2016). *Türkiye Bankacılık Sistemindeki Mevduat Bankaları ile Katılım Bankalarının Finansal Oranlar Açısından Lojistik Regresyon Analizi ile Değerlendirilmesi.* (Yayımlanmamış Yüksek Lisans Tezi), Balıkesir Üniversitesi, Balıkesir.

Kılıç, İ., Saraçlı, S. & Kolıkısaoğlı, S. (2011). Sosyo-Ekonomik Göstergeler Bakımından İllerin Bölgesel Bazda Benzerliklerinin Çok Değişkenli Analizler ile İncelenmesi. *İstatistikçiler Dergisi; İstatistik ve Aktüerya*, 4(2), 57-68.

Köklü, A. (1967). *Para ve Kredi*. Sevinç Matbaa, Ankara

Meydan. B. (2016). *Bankacılık ve Finans verileri doğrultusunda Türkiye'deki illerin sıralanması.* (Yayımlanmamış Yüksek Lisans Tezi), Hacettepe Üniversitesi, Ankara.

Önder, Z. & Özyıldırım, S. (2010). Banka Kredilerinin Bölgesel Büyümeye Etkileri. *ODTÜ Gelişme Dergisi*, 37(3), 267-292.

Özdamar, K. (2013). *Paket Programlar ile İstatistiksel Veri Analizi*. (Cilt 2, 9th ed.). Nisan Kitabevi, Ankara.

Sarıgül, H. (2014). Kümeleme Analizi ile İllerin Bankacılık Hizmetlerine Erişim ve Kullanım Göstergelerinin Karşılaştırılması. *Bankacılık Dergisi*, 25(89), 41-62.

Sosyo-Ekonomik Gelişmişlik Sıralaması Araştırmaları (SEGE). (2017). "*İllerin ve bölgelerin sosyo-ekonomik gelişmişlik sıralaması araştırması çalışma raporu*. T.C. Sanayi ve Teknoloji Bakanlığı Kalkınma Ajansları Genel Müdürlüğü. https://www.sanayi.gov.tr/merkez-birimi/b94224510b7b/sege/il-sege-raporlari.

Şener, O.H. (2019). *Banka Mevduatı ve Hukuki Niteliği* (2nd.Ed.). Yetkin Yayınları, Ankara.

Tatlıdil, H. (2002). *Uygulamalı Çok Değişkenli İstatistiksel Analiz*. Akademi Matbaası, Ankara.

Tekin, B. & Temelli. F. (2020a). Türkiye'deki İllerin Kredi Kullanım Düzeyleri Açısından Kümelenmesi: K- Ortalamalar Yöntemi Yaklaşımı. *Kastamonu Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 22(1), 91-124. DOI: iibfdkastamonu.734428.

Tekin, B. & Temelli. F. (2020b). K-Ortalamalar Yöntemi ile Bankaların Sermaye Yeterlilik Rasyolarına Göre Kümelenmesi. *Kırıkkale University Journal of Social Sciences (KUJSS),* 10(1), 11-36.

URL-1: Taşkan, P. (2006). *İstatistiki Bölge Birimleri Sınıflandırması (İBBS*). TÜİK Sunumu, Ankara, 20-22 Aralık, ss. 1-31. Retrieved March 12, 2024, from https://www.slideserve.com/wilda/istatistiki-b-lge-birimleri-siniflandirmasi-ibbs

URL-2: Türkiye Bankalar Birliği – Homepage (n.d.) Retrieved April 4, 2024, from https://verisistemi.tbb.org.tr/index.php?/tbb/report_bolgeler

URL-3: Türkiye İstatistik Kurumu – Homepage (n.d.) Retrieved April 4, 2024 from https://data.tuik.gov.tr/Search/Search?text=n%C3%BCfus

Yaman Yılmaz, C. (2020). Türkiye'deki İllerin Bankacılık Faaliyetleri Açısından Çok Değişkenli İstatistiksel Yöntemler ile İncelenmesi. *Alanya Akademik Bakış Dergisi*, 4(2), 471-493. https://doi.org/10.29023/alanyaakademik.685945

Yılmaz, Z. & Uzgören, E. (2013). Türkiye'de İllerin Temel Bankacılık Faaliyetleri Yönünden Kümeleme Analizi Yöntemiyle Sınıflandırılması. *Dumlupınar Üniversitesi Sosyal Bilimler Dergisi,* EYİ 2013, Özel Sayısı, 535-554.

Appendix

Appendix Table 1. Classification of provinces in Turkey according to per capita bank deposits using the complete linkage clustering method

Number of clusters (k)	No	Provinces in the cluster	Number of provinces
4	1	(34)	1
	2	(06)	1
	3	(35), (07), (16), (48)	4
	4	(41), (26), (59), (20), (54), (77), (14), (42), (09), (32), (38), (39), (11), (17), (22), (78), (45), (10), (33), (61), (01), (67), (64), (27), (55), (15), (71), (81), (70), (53), (43), (05), (31), (50), (03), (23), (40), (44), (58), (74), (24), (37), (08), (19), (68), (57), (28), (80), (18), (60), (51), (46), (62), (52), (25), (79), (66), (29), (69), (02), (75), (21), (36), (76), (12), (72), (63), (47), (56), (13), (65), (30), (49), (04), (73)	75
5	1	(34)	1
	2	(06)	1
	3	(35), (07), (16), (48)	4
	4	(41), (26), (59), (20), (77), (09), (38), (39), (17), (22), (10), (33), (61), (01), (67), (64), (50), (40), (68), (62)	20
	5	(54), (14), (42), (32), (11), (78), (45), (27), (55), (15), (71), (81), (70), (53), (43), (05), (31), (03), (23), (44), (58), (74), (24), (37), (08), (19), (57), (28), (80), (18), (60), (51), (46), (52), (25), (79), (66), (29), (69), (02), (75), (21), (36), (76), (12), (72), (63), (47), (56), (13), (65), (30), (49), (04), (73)	55
6	1	(34)	1
	2	(06)	1
	3	(35), (07), (16), (48)	4
	4	(41), (20), (38), (64), (50), (40), (68), (62),	8
	5	(26), (59), (77), (09), (39), (17), (22), (10), (33), (61), (01), (67)	12
	6	(54), (14), (42), (32), (11), (78), (45), (27), (55), (15), (71), (81), (70), (53), (43), (05), (31), (03), (23), (44), (58), (74), (24), (37), (08), (19), (57), (28), (80), (18), (60), (51), (46), (52), (25), (79), (66), (29), (69), (02), (75), (21), (36), (76), (12), (72), (63), (47), (56), (13), (65), (30), (49), (04), (73).	55

Appendix Table 2. Classification of provinces in Turkey according to per capita bank loans using the complete linkage clustering method

Number of clusters (k)	No	Provinces in the cluster	Number of provinces
4	1	(34), (06)	2
	2	(35), (07), (16), (20), (27)	5
	3	(41), (26), (48), (59), (42), (38), (39), (17), (22), (10), (61), (01), (53), (19)	14
	4	(54), (77), (14), (09), (32), (11), (78), (45), (33), (67), (64), (55), (15), (71), (81), (70), (43), (05), (31), (50), (03), (23), (40), (44), (58), (74), (24), (37), (08), (68), (57), (28), (80), (18), (60), (51), (46), (62), (52), (25), (79), (66), (29), (69), (02), (75), (21), (36), (76), (12), (72), (63), (47), (56), (13), (65), (30), (49), (04), (73)	60

Appendix Table 2. (Continue)

Number of clusters (k)	No	Provinces in the cluster	Number of provinces
5	1	(34), (06)	2
	2	(35), (07), (16), (20), (27)	5
	3	(41), (26), (48), (42), (38), (61), (01), (53), (19)	9
	4	(59), (39), (17), (22), (10)	5
	5	(54), (77), (14), (09), (32), (11), (78), (45), (33), (67), (64), (55), (15), (71), (81), (70), (43), (05), (31), (50), (03), (23), (40), (44), (58), (74), (24), (37), (08), (68), (57), (28), (80), (18), (60), (51), (46), (62), (52), (25), (79), (66), (29), (69), (02), (75), (21), (36), (76), (12), (72), (63), (47), (56), (13), (65), (30), (49), (04), (73)	60
6	1	(34), (06)	2
	2	(35), (07), (16), (20), (27)	5
	3	(41), (26), (48), (42), (38), (61), (01), (53), (19)	9
	4	(59), (39), (17), (22), (10),	5
	5	(54), (77), (14), (09), (32), (11), (78), (45), (33), (67), (64), (55), (15), (71), (81), (70), (43), (05), (31), (50), (03), (23), (40), (44), (58), (74), (37), (08), (68), (57), (28), (80), (18), (60), (51), (46), (62), (52), (25), (12)	40
	6	(24), (79), (66), (29), (69), (02), (75), (21), (36), (76), (72), (63), (47), (56), (13), (65), (30), (49), (04), (73)	20

Appendix Table 3. Classification of Turkish provinces based on per capita bank deposits using the median clustering method

Number of clusters (k)	No	Provinces in the cluster	Number of provinces
4	1	(34)	1
	2	(06)	1
	3	(35), (07), (16), (48)	4
	4	(41), (26), (59), (20), (54), (77), (14), (42), (09), (32), (38), (39), (11), (17), (22), (78), (45), (10), (33), (61), (01), (67), (64), (27), (55), (15), (71), (81), (70), (53), (43), (05), (31), (50), (03), (23), (40), (44), (58), (74), (24), (37), (08), (19), (68), (57), (28), (80), (18), (60), (51), (46), (62), (52), (25), (79), (66), (29), (69), (02), (75), (21), (36), (76), (12), (72), (63), (47), (56), (13), (65), (30), (49), (04), (73)	75
5	1	(34)	1
	2	(06)	1
	3	(35), (07), (16), (48)	4
	4	(41), (26), (59), (20), (54), (77), (14), (42), (09), (32), (38), (39), (11), (17), (22), (78), (45), (10), (33), (61), (01), (67), (64), (27), (55), (15), (70), (53), (31), (50), (03), (40), (58), (74), (24), (37), (08), (19), (68), (57), (28), (62), (52), (66)	44
	5	(71), (81), (43), (05), (23), (44), (80), (18), (60), (51), (46), (25), (79), (29), (69), (02), (75), (21), (36), (76), (12), (72), (63), (47), (56), (13), (65), (30), (49), (04), (73).	31

Appendix Table 3. (Continue)

Number of clusters (k)	No	Provinces in the cluster	Number of provinces
6	1	(34)	1
	2	(06)	1
	3	(35), (07), (16), (48)	4
	4	(41), (20)	2
	5	(26), (59), (54), (77), (14), (42), (09), (32), (38), (39), (11), (17), (22), (78), (45), (10), (33), (61), (01), (67), (64), (27), (55), (15), (70), (53), (31), (50), (03), (40), (58), (74), (24), (37), (08), (19), (68), (57), (28), (62), (52), (66)	42
	6	(71), (81), (43), (05), (23), (44), (80), (18), (60), (51), (46), (25), (79), (29), (69), (02), (75), (21), (36), (76), (12), (72), (63), (47), (56), (13), (65), (30), (49), (04), (73).	31

Appendix Table 4. Classification of Turkish provinces according to bank loans per capita using the median clustering method

Number of clusters(k)	No	Provinces in the cluster	Number of provinces
4	1	(34), (06)	2
	2	(35), (41), (07), (16), (26), (48), (20), (42), (38), (61), (01), (27), (53), (19)	14
	3	(59), (39), (17), (22), (10)	5
4	4	(54), (77), (14), (09), (32), (11), (78), (45), (33), (67), (64), (55), (15), (71), (81), (70), (43), (05), (31), (50), (03), (23), (40), (44), (58), (74), (24), (37), (08), (68), (57), (28), (80), (18), (60), (51), (46), (62), (52), (25), (79), (66), (29), (69), (02), (75), (21), (36), (76), (12), (72), (63), (47), (56), (13), (65), (30), (49), (04), (73)	60
5	1	(34), (06)	2
	2	(35), (07), (16), (20), (27)	5
	3	(41), (26), (48), (42), (38), (61), (01), (53), (19)	9
	4	(59), (39), (17), (22), (10)	5
	5	(54), (77), (14), (09), (32), (11), (78), (45), (33), (67), (64), (55), (15), (71), (81), (70), (43), (05), (31), (50), (03), (23), (40), (44), (58), (74), (24), (37), (08), (68), (57), (28), (80), (18), (60), (51), (46), (62), (52), (25), (79), (66), (29), (69), (02), (75), (21), (36), (76), (12), (72), (63), (47), (56), (13), (65), (30), (49), (04), (73).	60
6	1	(34)	1
	2	(06)	1
	3	(35), (07), (16), (20), (27)	5
	4	(41), (26), (48), (42), (38), (61), (01), (53), (19)	9

Appendix Table 5. Classification of provinces in Turkey according to per capita bank deposits using the ward clustering method

Number of clusters (k)	No	Provinces in the cluster	Number of provinces
4	1	(34), (06)	2
	2	(35), (07), (16), (48)	4
	3	(41), (26), (59), (20), (77), (09), (38), (39), (17), (22), (10), (33), (61), (01), (67), (64), (50), (40), (68), (62)	20

Appendix Table 5. (Continue)

Number of clusters (k)	No	Provinces in the cluster	Number of provinces
4	4	(54), (14), (42), (32), (11), (78), (45), (27), (55), (15), (71), (81), (70), (53), (43), (05), (31), (03), (23), (44), (58), (74), (24), (37), (08), (19), (57), (28), (80), (18), (60), (51), (46), (52), (25), (79), (66), (29), (69), (02), (75), (21), (36), (76), (12), (72), (63), (47), (56), (13), (65), (30), (49), (04), (73)	55
5	1	(34), (06)	2
	2	(35), (07), (16), (48)	4
	3	(41), (26), (59), (20), (77), (09), (38), (39), (17), (22), (10), (33), (61), (01), (67), (64), (50), (40), (68), (62)	20
	4	(54), (14), (42), (32), (11), (78), (45), (27), (55), (15), (71), (81), (70), (53), (43), (05), (31), (03), (23), (44), (58), (74), (24), (37), (08), (19), (57), (28), (18), (60), (51), (46), (52), (66), (29), (76),	36
	5	(80), (25), (79), (69), (02), (75), (21), (36), (12), (72), (63), (47), (56), (13), (65), (30), (49), (04), (73).	19
6	1	(34)	1
	2	(06)	1
	3	(35), (07), (16), (48)	4
	4	(41), (26), (59), (20), (77), (09), (38), (39), (17), (22), (10), (33), (61), (01), (67), (64), (50), (40), (68), (62)	20
	5	(54), (14), (42), (32), (11), (78), (45), (27), (55), (15), (71), (81), (70), (53), (43), (05), (31), (03), (23), (44), (58), (74), (24), (37), (08), (19), (57), (28), (18), (60), (51), (46), (52), (66), (29), (76)	36
	6	(80), (25), (79), (69), (02), (75), (21), (36), (12), (72), (63), (47), (56), (13), (65), (30), (49), (04), (73).	19

Appendix Table 6. Classification of provinces in Turkey according to per capita bank loans using the ward clustering method

Number of clusters (k)	No	Provinces in the cluster	Number of provinces
4	1	(34), (06)	2
	2	(35), (07), (16), (20), (27)	5
	3	(41), (26), (48), (59), (42), (38), (39), (17), (22), (10), (61), (01), (53), (19)	14
	4	(54), (77), (14), (09), (32), (11), (78), (45), (33), (67), (64), (55), (15), (71), (81), (70), (43), (05), (31), (50), (03), (23), (40), (44), (58), (74), (24), (37), (08), (68), (57), (28), (80), (18), (60), (51), (46), (62), (52), (25), (79), (66), (29), (69), (02), (75), (21), (36), (76), (12), (72), (63), (47), (56), (13), (65), (30), (49), (04), (73)	60
5	1	(34), (06)	2
	2	(35), (07), (16), (20), (27)	5
	3	(41), (26), (48), (42), (38), (61), (01), (53), (19)	9
	4	(59), (39), (17), (22), (10)	5

Appendix Table 6. (Continue)

Number of clusters (k)	No	Provinces in the cluster	Number of provinces
5	5	(54), (77), (14), (09), (32), (11), (78), (45), (33), (67), (64), (55), (15), (71), (81), (70), (43), (05), (31), (50), (03), (23), (40), (44), (58), (74), (24), (37), (08), (68), (57), (28), (80), (18), (60), (51), (46), (62), (52), (25), (79), (66), (29), (69), (02), (75), (21), (36), (76), (12), (72), (63), (47), (56), (13), (65), (30), (49), (04), (73)	60
6	1	(34), (06)	2
	2	(35), (07), (16), (20), (27)	5
	3	(41), (26), (48), (42), (38), (61), (01), (53), (19)	9
	4	(59), (39), (17), (22), (10)	5
	5	(54), (77), (14), (09), (32), (11), (78), (45), (33), (67), (64), (55), (15), (81), (70), (43), (05), (31), (50), (03), (23), (40), (44), (74), (37), (08), (68), (57), (28), (80), (18), (60), (51), (46), (62), (52), (25), (12)	38
	6	(71), (58), (24), (79), (66), (29), (69), (02), (75), (21), (36), (76), (72), (63), (47), (56), (13), (65), (30), (49), (04), (73).	22

Appendix Table 7. Classification of provinces in Turkey according to per capita bank deposits using the k-means clustering method

Number of clusters (k)	No	Provinces in the cluster	Number of provinces
4	1	(34)	1
	2	(06)	1
	3	(35), (41), (07), (16), (48), (20)	6
	4	(26), (59), (54), (77), (14), (42), (09), (32), (38), (39), (11), (17), (22), (78), (45), (10), (33), (61), (01), (67), (64), (27), (55), (15), (71), (81), (70), (53), (43), (05), (31), (50), (03), (23), (40), (44), (58), (74), (24), (37), (08), (19), (68), (57), (28), (80), (18), (60), (51), (46), (62), (52), (25), (79), (66), (29), (69), (02), (75), (21), (36), (76), (12), (72), (63), (47), (56), (13), (65), (30), (49), (04), (73)	73
5	1	(34)	1
	2	(06)	1
	3	(41), (26), (59), (20), (77), (09), (32), (38), (39), (17), (22), (10), (33), (61), (01), (67), (64), (15), (70), (50), (40), (74), (68), (57), (28), (62), (66),	27
	4	(54), (14), (42), (11), (78), (45), (27), (55), (71), (81), (53), (43), (05), (31), (03), (23), (44), (58), (24), (37), (08), (19), (80), (18), (60), (51), (46), (52), (25), (79), (29), (69), (02), (75), (21), (36), (76), (12), (72), (63), (47), (56), (13), (65), (30), (49), (04), (73)	48
	5	(35), (07), (16), (48)	4
6	1	(34)	1
	2	(06)	1
	3	(35), (07), (16), (48)	4
	4	(41), (20), (38), (64), (50), (40), (68), (62)	8
	5	(26), (59), (77), (09), (32), (39), (17), (22), (10), (33), (61), (01), (67), (55), (15), (70), (31), (03), (74), (19), (57), (28), (66),	23
	6	(54), (14), (42), (11), (78), (45), (27), (71), (81), (53), (43), (05), (23), (44), (58), (24), (37), (08), (80), (18), (60), (51), (46), (52), (25), (79), (29), (69), (02), (75), (21), (36), (76), (12), (72), (63), (47), (56), (13), (65), (30), (49), (04), (73).	44

Appendix Table 8. Classification of provinces in Turkey according to bank loans per capita using the k-means clustering method

Number of clusters (k)	No	Provinces in the cluster	Number of provinces
4	1	(34), (06)	2
	2	(59), (39), (17), (22), (10),	5
	3	(35), (41), (07), (16), (48), (20), (27),	7
	4	(26), (54), (77), (14), (42), (09), (32), (38), (11), (78), (45), (33), (61), (01), (67), (64), (55), (15), (71), (81), (70), (53), (43), (05), (31), (50), (03), (23), (40), (44), (58), (74), (24), (37), (08), (19), (68), (57), (28), (80), (18), (60), (51), (46), (62), (52), (25), (79), (66), (29), (69), (02), (75), (21), (36), (76), (12), (72), (63), (47), (56), (13), (65), (30), (49), (04), (73)	67
5	1	(34), (06)	2
	2	(41), (26), (48), (42), (09), (38), (78), (45), (33), (61), (01), (64), (53), (31), (19),	15
	3	(59), (39), (17), (22), (10)	5
	4	(54), (77), (14), (32), (11), (67), (55), (15), (71), (81), (70), (43), (05), (50), (03), (23), (40), (44), (58), (74), (24), (37), (08), (68), (57), (28), (80), (18), (60), (51), (46), (62), (52), (25), (79), (66), (29), (69), (02), (75), (21), (36), (76), (12), (72), (63), (47), (56), (13), (65), (30), (49), (04), (73)	54
	5	(35), (07), (16), (20), (27)	5
6	1	(34), (06)	2
	2	(26), (48), (42), (09), (38), (78), (45), (33), (61), (01), (64), (53), (31), (19),	14
	3	(35), (41), (27),	3
	4	(54), (77), (14), (32), (11), (67), (55), (15), (71), (81), (70), (43), (05), (50), (03), (23), (40), (44), (58), (74), (24), (37), (08), (68), (57), (28), (80), (18), (60), (51), (46), (62), (52), (25), (79), (66), (29), (69), (02), (75), (21), (36), (76), (12), (72), (63), (47), (56), (13), (65), (30), (49), (04), (73).	54
	5	(07), (16), (20),	3
	6	(59), (39), (17), (22), (10)	5