Chapter 18
Using Clustering of Panel Data to Examine Housing Demand of Expatriate Turks and Foreigners: An Application of k-prototype Algorithm

Özlem Akay, Cahit Çelik, and Gülsen Kıral

Abstract In recent years, development in the Turkey's economic structure and implementation of the "economic and financial policy" shows significant effects on the housing sector. For this reason, according to economic factors affecting housing demand, it is important to find cities and countries showing common characteristics and to examine the developments in the housing market. In this study, it was aimed to determine the common characteristics of the provinces that preferred by the Turks living abroad and foreigners and of the countries that prefer housing demand from these provinces in Turkey. Two different panel data sets were created for expatriate Turks and foreigners who prefer the demand for housing in Turkey. Cluster analysis was performed to these data sets using k-prototypes algorithm Cluster validity indexes were calculated to determine the appropriate number of clusters for both data sets. According to the findings, while the optimal number of clusters for data set defined for the expatriate Turk preferring housing demand in Turkey was found to be six, the optimal number of clusters for the data set defined for foreign countries preferring housing demand from Turkey was found to be nine.

Keywords Housing demand · Clustering of panel data · k-prototypes algorithm · Turkey

Ö. Akay (✉)
Department of Statistics, Cukurova University, 01330 Adana, Turkey
e-mail: oakay@cu.edu.tr

C. Çelik · G. Kıral
Department of Econometrics, Cukurova University, 01330 Adana, Turkey
e-mail: cahit.celik@outlook.com

G. Kıral
e-mail: gkiral@cu.edu.tr

D. Uzun Ozsahin et al. (eds.), *Application of Multi-Criteria Decision Analysis in Environmental and Civil Engineering*, Professional Practice in Earth Sciences,
https://doi.org/10.1007/978-3-030-64765-0_18

18.1 Introduction

The sensitivity of the construction and housing sector to the general economic conditions is different in each country. The construction and housing sector in the global market has extensive experience and potential. The construction sector is the locomotive sector since it affects more than 250 sub-sectors connected to it. The construction sector, which is based on domestic capital, is also absorbing unemployment due to create employment related to hundreds of occupations. The employment potential created for production by Turkish construction and sub-sector components provides important advantages for the national economy. In addition, housing not only contains economic and demographic factors, but also has socioeconomic, sociodemographic and sociopsychological qualities. According to TurkStat data, the expatriate Turks, which turned the rise of foreign exchange into an opportunity to invest in real estate in their countries, accelerated the purchase of immovable properties, especially in Izmir and Istanbul. Turkish real estate purchases increased by 23% compared to the same period of 2016. After the amendment of the reciprocity law, there was an explosion in selling property to foreigners.

Middle East countries and other countries buying heavily housing from Turkey prefer green and cool regions such as Izmit and Trabzon as well as Istanbul, Antalya, Izmir and Yalova. According to TurkStat data, existing home sales to foreign countries from Turkey, the record levels reached 40 thousand in 2018. This increase in housing sales was influenced by the increase in exchange rates of 2018, a value-added tax (VAT) exemption for non-Turkish citizens, the reduction of the value to be paid for foreigners to obtain Turkish citizenship from 1 million dollars to 250 thousand dollars, and international promotion and fair activities. Turkish citizens living in Germany and foreigners wishing to acquire immovable property in Turkey is able to land transactions with Berlin Land Registry and Cadastre Office established in Consulate General of Berlin by General Directorate of Land Registry and Cadastre. In addition to procurement transactions, deed registration sample, mortgage, donation transactions and title deed registration, exchange, transfer, applications and map related to cadastral data can be done in this agency. The land registry and cadastre offices planned to be installed in any two countries, Denmark, Austria, Germany, Belgium, Greece, Russia, Britain, the Netherlands, Norway, Qatar and Turkey. While determining these countries, the number of Turkish citizens living in the country, the number of foreigners who acquire real estate, high potential countries and economic and historical relations are taken into consideration.

In this study, panel data sets were created by obtaining economic indexes data of the expatriate Turks demanding house in Turkey for monthly in 2013–2016 and of the foreign countries demand for housing in Turkey for monthly in 2015–2016. The data used in the analysis were taken from TURKSTAT (Turkey Statistical Institute). When households living outside Turkey are demanding house in Turkey, according to their expectation, cluster analysis was performed using the k-prototype algorithm to find the similar characteristics of preferred provinces and the countries they live in. R program was used for analysis.

The study consists of five section. In the first section of the study, the importance of housing demand in terms of Turkey's economy, Turks living abroad and foreign countries interest in housing demand from Turkey are described. The second part consists of extensive literature research. In the third part of the study, clustering of panel data, k-prototype algorithm and cluster validity indexes are disclosed to be used in the application phase. The fourth part of the research consists of the application Clustering of panel data were made by using k-prototype algorithm for both expatriate Turks and foreign countries with the factors affecting housing demand and interpreted in detail. In the last section of the study, the results are given. This study carries original research quality since clustering of panel data were made by using k-prototypes algorithm in order to determine interest of the individuals living outside Turkey who are demanding house in Turkey.

18.2 Literature Review

Carliner (1973) offered new evidence based on better data than has been available to earlier researchers. Using a four-year panel study which followed up movers, permanent income was defined and calculated in two ways. Then regressions were run on house value and rent on permanent income, price, age, race and sex of head. The results obtained are robust with respect to the definition of permanent income, and considerably lower than results from time series analysis or from cross-section studies that relied on grouped rather than individual data.

Hanushek and Quigley (1980) focused on household price sensitivity. However other changes such as income or family size in household conditions clearly affect housing consumption. Actually, given limited longitudinal data, information about other demand adjustments provides valuable information about consumption dynamics, given limited longitudinal data. Price changes can be seen as one of the various external effects on housing demand.

Goodman (1988) addressesed the determination of permanent income, housing price,housing demand and tenure choice. Full housing demand elasticities incorporate the interactive effects among the four stages of the model. Price and income have major effects in the tenure choice equation. Socio demographic variables, such asage, have complex effects that may be lost in simpler forms of estimation.

Bajari and Kahn (2005) presented a three-stage, nonparametric estimation procedure to recover willingness to pay for housing attributes. Firstly, they estimated a nonparametric hedonic home price function. Secondly, they recovered each consumer's taste parameters for product characteristics using first-order conditions for utility maximization. Finally, they estimated the distribution of household tastes as a function of household demographics.

Nunes and Serrasqueiro (2007) showed that internal and external financing are not perfect substitutes using panel data for the period 1999–2003, not corroborating the theorem of Modigliani and Miller.

Nunes et al. (2009) studied the profitability determinants of Portuguese service industries based on various panel models.

Davis et al. (2017) stated that households holding an FHA mortgage increased the value of the housing they purchased by approximately 2.5% using the financing of Fannie Mae and Freddie Mac. After converting the premium discount to an equivalent decrease in the mortgage rate, conditional on purchasing a home with roughly 3.4 estimates means semi-elasticity of the value of the housing purchased on the mortgage rate.

Nguyen and Nordman (2018) shaded light on the links between households' and entrepreneurs' social networks and business performance by using a unique panel of household businesses for Vietnam.

Do et al. (2019) examined the determinants of livestock assets with panel data from Vietnam. They suggested that authorising rural households to better cope with shocks contributes to reducing rural poverty and to developing livestock. Ahmad et al. (2018) investigated the determinants of housing demand in urban areas of Pakistan. Empirical analysis was performed using the 2004–05 and 2010–11 Pakistan Social and Standard of Living Measurement (PSLM) survey. The hedonic price model was used to estimate housing prices. Heckman's two-stage selection procedure is used to control the bias of selectivity between duty term selection and the amount of housing services requested. Empirical analysis shows that house prices and revenues (temporary and permanent) play an important role in determining the demand for housing units.

Huarng et al. (2019) analyzed housing demand by using Google Trends' big data as a proxy. They use to estimate a qualitative method (fuzzy set/ Qualitative Comparative Analysis, fsQCA) instead of a quantitative method. According to empirical results, although the size of the data set is small, fsQCA successfully predicts seasonal time series.

Zheng et al. (2018) estimated the income elasticity of demand for private rental housing using micro data between 1996 and 2011 from four waves of four Hong Kong census data. In order to isolate permanent and temporary income at the household level they adopted a permanent income model. They used the Heckman two-step procedure to correct selection bias and used the quantile regression (QR) approach to investigate the heterogeneity of demand elasticities between different levels of housing expenditure. Empirical results show that permanent income elasticities fall within the range of 0.536–0.698 and that temporary income shock has a positive and significant impact on rental housing demand.

Liu (2019) examined the theoretical relationship between income and home prices by using the user cost equilibrium condition. Empirically, the short-term and long-term dynamics of this relationship studied from 1991 to 2015 for 25 years in the state of New South Wales, Australia, using data for 144 Local Government Areas (LGAs). He estimated to be 1.07 the income elasticity of housing prices for the government with multi-factor panel data models and cointegration analysis.

Çelik and Kiral (2018a) applied balanced and unbalanced panel data analysis and clustering analysis methods to factors that affect housing sales in provinces of Turkey and examined significant variables by hierarchical clustering method. They also supported the study with SWOT analysis. Çelik and Kıral (2018b) used clustering of panel data and SWOT analysis to examined the socio-economic factors affecting housing sales of the provinces of Turkey in the 2008–2015 process. They determined factors affecting housing sales in provinces that exhibit similar characteristics in housing demand. The results obtained from the study showed that urbanization rate, the ratio of deposit interest rate, average household income, number of household's automobiles, stock market Istanbul 100 index, housing loan interest rateangross return rate of housing, were significant in for housing demand. Akay and Yüksel (2018) presented that the mixed panel dataset is clustered by agglomerative hierarchical algorithms based on Gower's distance and by k-prototypes . Akay and Yüksel (2019) suggested a new distance for clustering of the mixed variable panel data set containing invariant time binary variable, without performing variable conversion to avoid information loss.

18.3 Cluster Analysis of Panel Data with k-Prototype Algorithm

Panel data refer to two-dimensional data which are obtained in time series and cross section at the same time, and that means taking multiple cross sections on time series, and selecting the sample observations on cross sections at the same time (Hou and Ai 2015). The poobility of the different topics in the data is one of the important issues in the panel data. If the parameters in the regression can be considered homogeneous between different subjects, different subjects can be brought together. However, the normal situation is that subjects cannot be pooled due to high heterogeneity. Some recent studies investigated the "partial poolability" by clustering subjects into different groups so that subjects in the same cluster have homogeneous parameters (Lu and Huang 2011).

Bonzo and Hermosilla (2002) applied probability link function to advance the algorithm of the cluster, thus the cluster analysis could be effectively applied to the analysis panel data. In this study, we choose k prototype algorithm to explain the cluster analysis process of multivariable panel data.This algorithm was proposed by Huang (1998). It is straightforward to integrate the k-modes and k-means algorithms into the k-prototypes algorithm used to cluster the mixed-type objects. Since frequently encountered objects in real world databases are mixed-type objects, the k-prototypes algorithm is practically more useful. The cost function is used in conjunction with a partitioned clustering algorithm. The cost function handles mixed datasets and computes the distance between a data point and a centre of cluster in terms of two distance values—one for the numeric attributes and the other for the categorical attributes.The objective of k-prototype is to group the dataset X into k clusters by minimizing the cost function,

$$E = \sum_{l=1}^{k} \sum_{i=1}^{n} y_{il} d(X_i, Q_l) \tag{18.1}$$

Here, $Q_l = [q_{l1}, q_{l2}, \ldots, q_{lm}]$ is the representative vector or prototype for cluster l, and y_{il} is an element of a partition matrix Y_{nxl}. $d(X_i, Q_l)$ is the dissimilarity measure defined as follows:

$$d(X_i, Q_l) = \sum_{j=1}^{p} \sum_{t=1}^{T} \left(X^r_{ij}(t) - q^r_{lj}\right)^2 + \mu_l \sum_{j=p+1}^{m} \delta\left(X^c_{ij}(t) - q^c_{lj}\right) \tag{18.2}$$

where $\delta(p, q) = 0$ for $p = q$, and $\delta(p, q) = 1$ for $p \neq q$; $X^r_{ij}(t)\left(X^c_{ij}(t)\right)$ is the value of the jth numeric (categorical) attribute at the time t for the data object i; $q^r_{lj}\left(q^c_{lj}\right)$ is the prototype of the jth numeric (categorical) attribute in the cluster l;μ_l is a weight for categorical attributes in the cluster l (Ji et al. 2012). The process of the k-prototype algorithm is described as follows:

The process of the K-prototype algorithm is defined as follows:

Step 1. Randomly select k data objects from the dataset X as the initial prototype of the sets.

Step 2. For each data object in X, assign it to the cluster whose prototype is closest to that data object in terms of Eq. (18.2). After each assignment, update the prototype of the cluster.

Step 3. After all data objects have been assigned to a cluster, recalculate the similarity of the data objects with the existing prototypes. If a data object is found to belong to another cluster rather than the closest prototype, reassign that data object to that cluster and update the prototypes of both clusters.

Step 4. After the full circle test of X, terminate the algorithm if no data object has changed the sets, or else repeat step 3 (Ji et al. 2013).

Different clustering algorithms often lead to different clusters of data, even for the same algorithm, the choice of different parameters or the order of presentation of data objects can greatly affect the final clusters. Therefore, effective assessment standards and criteria are critical to reassuring users of cluster results. For all that, these evaluations provide meaningful information on how many clusters are hidden in the data. Actually, the user is faced with the dilemma of selecting the number of clusters or partitions in the underlying data. Therefore, numerous indices have been proposed to determine the number of clusters in a data set (Charrad et al. 2012).

Some clustering validity indices are used to select the optimal number of clusters. These indices are The C-Index, Dunn index, Gamma index, Gplus index, McClain index, Ptbiserial index, Silhouette index and Tau index. The minimum values of the C-Index, Gplus and McClain index are used to indicate the optimal number of clusters. The maximum values of the Dunn, Gamma, Ptbiserial, Silhouette and Tau index are used to indicate the optimal number of clusters.

18.4 Application

This study aims to determine provinces in Turkey and countries showing common feature according to the expectations the Turks living abroad and foreign countries in housing demand. For this purpose, panel data set was formed, firstly, by monthly data for the years 2013–2016 for provinces where expatriates demand housing from Turkey (Antalya, Istanbul, Trabzon, Aydin, Bursa, Other Provinces, Mersin, Mugla, Yalova, Ankara, Izmir, Sakarya) and secondly by considering economic factor indices determining the demand for housing.

Economic factor indexes, European Currency, Bullion Gold, General Import, Economic Confidence Index, Consumer Confidence Index, Real Sector Confidence Index, Construction Sector Confidence Index, Employment Index, Employees in a Paid Job are determined as the shadow variable taking a value of 1 for investment purpose provinces and 0 for the others as described in detail in Table 18.1.

Subsequently, by considering the economic indices of foreign countries (Iraq, Kuwait, Russian Federation, Saudi Arabia, China, Qatar, Yemen, Belgium, United Arab Emirates, Egypt, Jordan, Libya, Ukraine, Netherlands, Norway, Afghanistan, Germany, Azerbaijan, Iran, United Kingdom, Sweden, Kazakhstan and Other Countries) which demand housing from Turkey, monthly panel data set has been formed for the years 2015–2016. Variables Monthly Number of House Sales to Foreigners by National Nation, European Currency, Gold Bullion, Economic Confidence Index, Consumer Confidence Index, Real Sector Confidence Index, Service Sector Confidence Index, Retail Trade Confidence Index are determined as the shadow variable that takes the value of 1 for investment purpose provinces and 0 for the others, and are described in detail in Tables 18.2 and 18.3.

Table 18.1 Turkey's economic data for expats in demand for housing in Turkey

KT	Housing demand (monthly number of residential sales in overseas settlements 48 months 2013–2016)
EURO	European currency/ Turkish Lira (exchange rate) monthly average 2013–2016
KALTIN	Gold ıngot selling gold price (TL/Gr) 2013–2016
D1	Dummy variable that takes value 1 for investment provinces and 0 for others
ITHALAT	General Imports (monthly foreign trade quantity indices by international standard trade classification 2013–2016)
EGE	Economic confidence Index (2013–2016)
TUGE	Consumer confidence Index (2013–2016)
REGE	Real sector confidence Index (2013–2016)
ISGE	Construction sector confidence Index (2013–2016)
ISTHEN	Employment index (seasonally and calendar adjusted index) trade and service indices and change rates (2013–2016)
UBIC	Paid employees (arriving citizens by general working status-residing in Turkey-2013–2016) TSI, citizen entry survey

Table 18.2 Located in housing demand from Turkey, foreign country economic data

KT	Monthly housing sales to foreigners by nationality (24 months 2015–2016)
EURO	European currency/Turkish Lira (exchange rate) average monthly 2015–2016
KALTIN	Gold ingot (selling gold price (TL/Gr) 2015–2016
D1	Dummy variable that takes value 1 for investment provinces and 0 for others
ITHALAT	General ımports (monthly foreign trade quantity ındices by ınternational standard trade classification 2013–2016)
EGE	Economic confidence ındex (2015–2016)
TUGE	Consumer confidence ındex (2015–2016)
REGE	Real sector confidence ındex (2015–2016)
HSGE	Service sector confidence ındex (2015–2016)
PTSGE	Retail trade sector confidence ındex (2015–2016)
D1	Shadow variable that takes value 1 for investment provinces and 0 for others

Table 18.3 The first ten countries that most housing purchases in the years 2015–2016 from Turkey

Country	2015 (Piece)	Country	2016 (Piece)	Rate of change (%)
England	4.552	Iraq	3.726	−15.5
Iraq	4.407	Britain	2.556	−43.8
Russia	2.377	Saudi Arabia	1.827	−20.3
Kuwait	2.299	Kuwait	1.749	−23.9
Saudi Arabia	2.292	Afghanistan	1.623	*
Germany	1.280	Germany	1.474	15.2
Azerbaijan	864	Russia	1.449	−39.0
United Arab Emirates	329	Azerbaijan	724	−16.2
Train	279	Lebanon	191	*
USA	263	TRNC	180	*

Source Eva real estate appraisal (2016)

Clustering analysis was performed to both panel data sets by using k-prototype algorithm. By using cluster validity indexes, the appropriate numbers of clusters were determined and similar units were obtained in housing demand. Cluster validity index values are given in Tables 18.4 and 18.5.

As shown in Table 18.4, for foreign resident data set requesting housing in Turkey, while cluster validity indices of Dunn (0.7467), Ptbiserial (0.7640), Silhouette (0.7733), Thai (0.5965) indices propose the optimal number of cluster as 2, Cindex (0.0111), Gamma (0.9348), Gplus (0.0093) index values recommend as 6. Considering that the number of clusters envisaged for the study will be 6, the number of appropriate clusters is taken as 6. The distribution of provinces in clusters for the number of clusters 6;

1st Cluster: Antalya, Istanbul.

Table 18.4 Cluster validity index values for non-resident data sets in the demand for housing in Turkey

Number of clusters	2	3	4	5	6	7
Cindex	0.0198	0.0224	0.0136	0.0140	**0.0111**	0.0154
Dunn	**0.7467**	0.3820	0.4915	0.3914	0.2132	0.6394
Gamma	0.8674	0.8484	0.8666	0.8831	**0.9348**	0.9289
Gplus	0.0298	0.0349	0.0209	0.0163	**0.0093**	0.0102
McClain	0.2517	0.2352	0.2102	0.2083	0.1808	**0.1380**
Ptbiserial	**0.7640**	0.5144	0.4748	0.4057	0.3336	0.3312
Silhouette	**0.7733**	0.6529	0.6529	0.5859	0.4877	0.5407
Tau	**0.5965**	0.5829	0.5788	0.3775	0.4480	0.3823

Table 18.5 Cluster validity index value for foreign countries in the demand for housing data sets from Turkey

Number of clusters	2	3	4	5	6	7	8	9	10
Cindex	0.0658	0.0182	0.0121	0.0085	0.0070	0.0059	0.0056	**0.0037**	0.0070
Dunn	0.0921	0.1698	0.1259	0.1474	0.1995	0.3361	0.2151	**0.3495**	0.2776
Gamma	0.8012	0.9445	0.9232	**0.94784**	0.9254	**0.94781**	0.9127	0.9333	0.9113
Gplus	0.0501	0.0122	0.0143	0.0130	0.0095	0.0097	0.0044	**0.0041**	0.0050
McClain	0.2668	0.1510	0.1291	0.1155	0.0991	0.0836	0.0840	0.0843	**0.0686**
Ptbiserial	**0.6200**	0.6026	0.5235	0.4782	0.4236	0.3740	0.3522	0.2992	0.3375
Silhouette	0.7064	**0.7707**	0.7471	0.6934	0.6793	0.6043	0.6441	0.6491	0.5725
Tau	0.5714	**0.6301**	0.5656	0.5035	0.4882	0.4508	0.4168	0.3919	0.3936

2nd Cluster: Trabzon.
3rd Cluster: Aydin, Bursa, Other Provinces, Mugla, Yalova.
4th Cluster: Ankara.
5th Cluster: Izmir.
6th Cluster: Sakarya.

as obtained. Comment of Clustering for External Residents;

Antalya is one of Turkey's major cities for tourism. Due to its comfortable living conditions, cultural and historical richness, nature and climate, it attracts the attention of investors from abroad and within the country. Antalya is one of the most immigration receiving provinces and its economy depends on tourism. However, expatriates invested their years of accumulation into residences in their respective cities. Now, implementation of attractive mortgage conditions in Turkey, which are widespread abroad, has caused the Turks to move to Istanbul.

TurkStat's research reveals that Turks living abroad invest in closed sites in Istanbul. Expatriates highlight Antalya and Istanbul in terms of housing investment, which provide both job opportunities and quality living standards, and enable them to be examined in the same cluster. Trabzon, Ankara, Izmir and Sakarya provinces are examined in a cluster by themselves in terms of housing demand. The most important reasons why the expatriate Turks prefer Trabzon in their housing investments are the cool and rainy climate and natural beauties, besides the green nature, housing prices are cheaper compared to provinces such as Istanbul and Antalya, and ease of transportation due to the fact that there is an airport in Trabzon. In addition, the most important reason for the expatriate Turks' housing investments in Ankara is the high potential of Ankara's industry, transportation, industry, education and technical, political, tourism, culture and arts.

On the other hand, Izmir is one of the most preferred cities in terms of housing investment and tourism for expatriate Turks living in Europe. The most important reason why expats prefer 2 + 1 or 3 + 1 house types is that they want to settle in Izmir after their retirement. Thanks to the junctions of the transportation roads belonging to the province of Sakarya and the developing industry, external and internal migrations go on. It is estimated that one million people will live in Sakarya, the shining star of Marmara Region.

Because thermal tourism is important for health, expats demand housing from these places. On the other hand, the real estate sector, which is faced with interest rates, rising costs and exchange rate pressures, Aydın, Bursa, Other Provinces, Mersin, Mugla and Yalova receives high demand from expatriates. Expatriate Turks working abroad and earning foreign exchange make their real estate investments in various cities of Turkey, where they come for holiday.

As shown in Table 18.5, for foreign resident housing demand data set in Turkey, while cluster validity indices of Silhouette (0.7707) and Tau (0.6301) index values propose the optimal number of cluster as three, Cindex (0.0037), Dunn (0.3495) and Gplus (0.0041) indices suggest the optimum number of clusters as 9. Considering that the number of clusters envisaged for the study will be 9, the appropriate number of clusters is taken as 9. The distribution of countries in clusters for 9 clusters;

1st Cluster: Other Countries, Iraq, Kuwait, Russian Federation, Saudi Arabia.
2nd Cluster: China, Qatar, Yemen.
3rd Cluster: Belgium.
4th Cluster: United Arab Emirates, Egypt.
5th Cluster: Jordan.
6th Cluster: Libya, Ukraine.
7th Cluster: Netherlands.
8th Cluster: Norway.
9th Cluster: Afghanistan, Germany, Azerbaijan, Iran, England, Sweden, Kazakhstan as obtained. Comment of Clustering for Foreign Countries;

The housing demand of the five countries in the first cluster and the seven countries in the ninth cluster was determined to be for investment and holiday purposes. In previous years, Afghans who were not seen among countries purchasing houses the most in Turkey, according to the 2016 TSI data, with the purchase of 1623 housings in Turkey are in fifth place in the top 10 most purchases.

Afghans demand housing especially from IstanbulWhen compared with the Germans housing demand of 2015, housing demand interest in Turkey has continued in 2016. In 2016, it ranks second in the list with 1474 units and 649,254 m^2 purchases. Another country that was.

on the list in the previous year and continued to increase in 2016 is Azerbaijan. Although the number of Azerbaijanis, which ranked ninth in the list in 2015, decreased in number, their purchases in square meters increased. Azerbaijanis ranked fifth with 280,247 m^2 purchases. It ranks eighth in the list of housing purchases.

According to a research by the Demir, Construction Management Board, border neighborhood of Turkey with Iran it is important from a strategic point of view. Thus, the Iranians are among the most residential area purchasing nations. According to Radikal (2008), the Turkish real estate market since 2003 and it was opened to foreign buyers of the situation, the British property buyers purchase housing from Turkey because of the low prices. Any housing on the Turkish coast can be purchased for as low as 35,000 lb. While the UK is ranked first in the list with 4,552 housing purchases in 2015, it is ranked second with 2,556 housing purchases in 2016.

The similarity of the countries in the second, third, fourth, fifth, sixth, seventh and eighth clusters stems from the fact that housing demand is for investment, accommodation and tourism purposes. In 2015, Iraq was the leading foreign investor with 4 228 units purchase followed by Saudi Arabia with 2704, Kuwait with 2130, Libya with 427, United Arab.

Emirates with 332, Qatar with 277, Egypt with 318, Jordan with 243 and Yemen with 231 houses. Middle Eastern countries mostly buy real estate from Istanbul, Izmit and Yalova provinces.

When we look at the year 2016, Iraqis have 3 thousand 36 houses. The monthly average purchase price of Iraqis, the leader of 2015, was 253. In 2016, Saudi Arabia citizens ranked second with 1886 and Kuwait citizens ranked third with 1744. Accordingly, in 2016, 18 thousand 391 houses were sold to foreigners. The Dutch have purchased 217, the Belgians 198.

18.5 Conclusion

In this study, clustering analysis with k-prototypes algorithm was applied to the economic factors affecting the sale of houses of provincial groups. The main purpose of the study is to classify the provinces and countries showing common feature according to the expectations of the Turks living abroad and foreign countries in the demand for housing in Turkey.

According to the analysis results, the cluster validity indices Cindex (0.0111), Gamma (0.9348) and Gplus (0.0093) as index values found for expatriates who demand housing from Turkey suggest the optimal number of clusters as 6. On the other hand, cluster validity indices Cindex (0.0037), Dunn (0.3495) and Gplus (0.0041) as the index found for foreigners who demand for housing in Turkey suggests the optimal number of clusters as 9. According to the result of clustering; The expectations of Turkish citizens living abroad and foreigners' housing demands have shown similarity.

Expatriates and foreigners mostly want the houses they buy for investment and holiday purposes to be both economical and durable. Expatriates, who prefer summer locations, hometowns and metropolitan cities in the purchase of real estate, revive the housing market. In addition, the high course of foreign exchange, like expatriates, stimulates foreign buyers in housing purchases. Experts in the housing sector state that the VAT advantage granted to foreigners is also granted to Turks residing abroad for more than 6 months with work and residence permits. As a result, the housing market experts predict that there will be significant activity in the Turkish housing market if foreign housing increases and private housing campaigns are provided to expatriate Turks and foreigners.

References

Ahmad A, Iqbal N, Siddiqui R (2018) Determinants of housing demand in urban areas of Pakistan: evidence from the PSLM. Pakistan Dev Rev 57:1–25

Akay Ö, Yüksel G (2018) Clustering the mixed panel dataset using Gower's distance and k-prototypes algorithms. Commun Stat Simul Comput 47:3031–3041

Akay Ö, Yüksel G (2019) Hierarchical clustering of mixed variable panel data based on new distance. Commun Stat Simul Comput 1–16

Bajari P, Kahn ME (2005) Estimating housing demand with an application to explaining racial segregation in cities. J Bus Econ Stat 2:20–33

Bonzo DC, Hermosilla AY (2002) Clustering panel data via perturbed adaptive simulated annealing and genetic algorithms. Adv Comp Syst 5:339–360

Carliner G (1973) Income elasticity of housing demand. Rev Econ Stat 55:528–532

Charrad M, Ghazzali N, Boiteau V, Niknafs A (2012) NbClust package: finding the relevant number of clusters in a dataset. J Stat Softw, Softw

Çelik C, Kıral G (2018a) Kümeleme yöntemiyle konut talebinin incelenmesi: Türkiye il grupları üzerine bir uygulama. Çukurova Üniversitesi Sosyal Bilimler Enstitüsü Dergisi 27:123–138

Çelik C, Kıral G (2018b) Yurtdışı yerleşiklerin ve dış ülkelerin konut taleplerini incelemede panel kümeleme analizi: Türkiye illeri örneği. Çukurova Üniversitesi İktisadi Ve İdari Bilimler Fakültesi Dergisi 22:305–324

Davis M, Oliner S, Peter T, Pinto E (2017) The impact of interest rates on house prices and housing demand: evidence from a quasi-experiment. Working Paper

Do TL, Nguyen TT, Grote U (2019) Livestock production, rural poverty, and perceived shocks: evidence from panel data for Vietnam. J Dev Stud 55:99–119

Goodman AC (1988) An econometric model of housing price, permanent income, tenure choice, and housing demand. J Urban Econ 23:327–353

Hanushek EA, Quigley JM (1980) What is the price elasticity of housing demand? Rev Econ Stat 62:449–454

Huarng KH, Yu THK, Rodriguez-Garcia M (2019) Qualitative analysis of housing demand using Google trends data. Econ Res Ekonomska Istraživanja 409–419

Hou F, Ai K (2015) The empirical research of relationship between consumption and income for Chinese urban residents. Open J Appl Sci 5:251

Huang ZX (1998) Extensions to the k-means algorithm for clustering large datasets with categorical values. Data Min Knowl Disc 2:283–304

Ji J, Pang W, Zhou C, Han X, Wang Z (2012) A fuzzy k-prototype clustering algorithm for mixed numeric and categorical data. Knowl-Based Syst 30:129–135

Ji J, Bai T, Zhou C, Ma C, Wang Z (2013) An improved k-prototypes clustering algorithm for mixed numeric and categorical data. Neurocomputing 120:590–596

Liu X (2019) The income elasticity of housing demand in New South Wales, Australia. Reg Sci Urban Econ 75:70–84

Lu H, Huang S (2011) Clustering panel data. In: SIAM international workshop on data mining held in conjunction with the 2011 SIAM international, conference on data mining. Arizona, USA: IEEE SMCS, pp 1–10

Nguyen CH, Nordman CJ (2018) Household entrepreneurship and social networks: panel data evidence from Vietnam. J Dev Stud 54:594–618

Nunes PJM, Serrasqueiro ZM (2007) Capital structure of Portuguese service industries: a panel data analysis. Serv Ind J 27:549–562

Nunes PJM, Serrasqueiro ZM, Sequeira TN (2009) Profitability in Portuguese service industries: a panel data approach. Serv Ind J 29:693–707

Zheng X, Xia Y, Hui EC, Zheng L (2018) Urban housing demand, permanent income and uncertainty: Microdata analysis of Hong Kong's rental market. Habitat Int 74:9–17