Time Series Analysis of Housing Demand: A Forecasting Model for Ankara, Turkey

SAFIYE TURGAY[1], CEYHUN KANBOLAT[1], ESMA SEDEF KARA[2]
[1]Department of Industrial Engineering,
Sakarya University,
54187, Esentepe Campus Serdivan-Sakarya,
TURKEY

[2]Rüstempaşa Mahallesi, İpekyolu Caddesi, No:120,
Sapanca (54600) Sakarya,
TURKEY

Abstract: - The property boom in Ankara surrounded by urban arrangements, calls for complicated forecasting approaches so that stakeholders can benefit from logical decision-making. The researchers apply up-to-date time series analysis methodology to forecast the housing demand in the area. It implicates the historical sales of housing and economic indicators combined with demographic factors are the sources that develop a comprehensive model of forecasting which allows to explore and track the intrinsic dynamics of the housing market. The methodology, in turn, is the application of cutting-edge statistical models and machine learning algorithms in the process of capturing the complex trend that is explicit in the time series data. In terms of our approach, we will include seasonality as well as trend components as well as those external factors, which affect the level of houses' demands. The study also analyzes the outcomes caused by economic shocks, public policies, and urban planning on housing market equilibrium. The study carried out demand forecasting concerning the sale of houses in Turkey which is supported by the data. The study is based on TURKSTAT numbers on the number of houses sold within the year 2021 (S.O.D) by Turkish provinces that cover Ankara province where the data is retrieved from. Considering the sales of houses in Ankara from 2014-2018 as a basis, this study intends to find a numerical forecasting model that is most suited to the observed dataset and thus, determine the number of houses sold in Ankara in the year 2019 using this particular method. Output from time series analysis provides the developers and investors with significant information by the way of anticipating market fluctuations, improving their investment strategies, and choosing the right policies according to the markets' needs. Moreover, an accurate model needs to be analyzed through serious validation techniques to identify its authenticity in its real-life examples. This research is, at the same time, an attempt to make progress in the field of demand forecasting in the real estate market as well as an attempt to provide stakeholders working in Ankara Province with a comprehensive guide while moving through a changing housing market. The utilization of technology and a careful investigation of relevant factors lends this study credibility as well as makes it a necessary literary component for those pursuing a deeper comprehension of housing demand in the region.

Key-Words: - Time Series Analysis, Real Estate Market, Ankara Province, Economic Indicators, Demographic Factors, Machine Learning Algorithms, Market Dynamics.

Received: March 13, 2024. Revised: September 3, 2024. Accepted: October 4, 2024. Available online: November 18, 2024.

1 Introduction

The real estate market in Ankara Province, which is certain to experience new unseen growth, lots of changes, and new dynamics, is at the turning point in this course. As the real estate demand leads the city to be shaped dominantly by the housing industry, accurate models of prediction will be inevitable for the stakeholders to deal with the intricacies and make informed decisions. The research approach that this paper embraces sets out to construct a comprehensive view of the housing demand as reflected by the time series analysis where a specific forecasting model adapted to Ankara's real estate market peculiarities is presented.

Ankara, the capital city of Türkiye, is now residing in a period of urbanization and population change, which definitude the residential property demand greatly. The complex game between economic factors, government policies, and social changes leads to the moving pattern of the housing market that requires the application of a multifacetical approach. Therefore, here the research is carried out in order to showcase the fastest growing area in forecasting models which provides clarity on housing demand over time. The focus of this undertaking is to produce a forecasting method that uses the time series regression analysis in the prediction of the demand for housing within Ankara Province. To integrate the historical housing sales data with the key economic indicators and demographic features, the model would aim to deliver exact and up-to-date predictions on the future trends of the real estate market, [1]. Moreover, the study attempts to ascertain the effect of external forces like refrains from governments and subjects of urban development on the housing needs. The approach that is used within the framework of this particular study comprises of these diverse components. A dynamic statistical model and machine learning algorithm are employed on the gathered data. Temporal patterns, trend factors, and exogenous variables that can affect the accuracy of the forecasting models are discovered through the examined processes. Integration of multisource data allows us to see the whole structure of the factors that determine housing demand in Ankara.

The research's significance is that it can provide real estate developers, investors, and decision-makers with possible decision points. Forecasting which comes with accuracy in the anticipated future of the market also helps in the optimization of existing investment strategies and implementation of targeted interventions. It is anticipated that the study's results will not only become a part of academic conversations but also act as a practical tool for decision-makers while managing the dynamic and served housing market setting of Ankara province.

The next parts of the paper discuss the data collection process, the time series analysis method applied, and the building of the forecasting model are explored in the following. Discussion of the results and findings is then undertaken in depth, succeeded by a comprehensive analysis of the effects for both real estate investors and realtors. Finally, the essay ends with the main findings and the proposals for future study of housing demand forecasting, a vibrant topic among the fields of urban planning.

2 Literature Survey

The volatile character of housing markets demands spotting the factors that influence house demand correctly. In the cultures of Turkey, especially Ankara district, with urbanization and population movement that is changing the directions of the real estate business, therefore, a survey was conducted via literature that will provide a broader view of how the forecasting model may work through the use of time series analysis, [2], [3].

Many research the statistical tools for the next steps of real estate trend modeling and forecasting, while time series analysis is one most widely used methods for that. For example ARIMA evaluation revealed the patterns in housing markets changed accordingly, [4]. Such fundamental role of statistical models shows to what degree they contribute to finding the recurring trends and seasonality within this type of data, [5].

A link between economic indicators and housing demand has been considered in several sources on real estate. The influence of economic factors like GDP growth and unemployment rates on the housing market dynamics cannot be overlooked, [6]. Recognizing these linkages is of course extremely important for the construction of a model that examines the economic picture influencing the housing requests in Ankara.

Demographic attributes take active part in the way housing demand is formed which the interaction of population growth, diversity of demographics, and dynamism of the housing market, [7]. Input from this research is considered in the integration of demographic characteristics as an element of the forecasting model, as it is a vital factor in understanding the ever-changing needs of Ankara's multifaceted society.

The most recent scientific breakthroughs in machine learning have made a diversity of techniques available for accurate real estate forecasting. Some of the researchers showed that machine learning algorithms and especially ensemble methods predict housing market trajectories with much great certainty, [8], [9], [10], [11], [12], [13]. Such literature is used to develop a combination of machine learning methods to the end that the forecasting model will not only be thorough but also adaptive to the changing housing market environment, [14], [15].

Policy initiatives and government development have about huge influence on the housing markets. In addition to policy changes and urban planning in forecasting models, [16], this issue needs to be addressed. This awareness is key in deciding to analyze how fiscal policies and the building of urban infrastructures have effects on housing demand tendencies in Ankara, [17].

Housing demand time series analysis for Ankara city, Turkey will be the subject of the current study, as the literature review lays a strong groundwork for it. This study is looking into the possibility of incorporating incorporating established methodologies and outcomes by developing a forecasting model based on economic indicators, demographic factors, and machine learning, [18], [19], [20]. The study combines various research from the literature to contribute a comprehensive and contextual relevant theme to the understanding and predicting housing demands in Ankara province, [21], [22], [23].

3 Methodology

The core of this research is based on a complete dataset that covers historical housing sale data for Ankara City in the last ten years. This dataset encompasses details about property deals and pricing patterns with relevant economic factors that took place during a given time period. It is also worth noting that we consider in our model metrical data, urban development records, and government policy changes to determine external factors affecting housing demand. This stage entails the imputation of missing data, accompanying extreme or outlying values, and correcting the formatting inconsistencies. Longitudinal data is in a temporal format, while results of various factors are normalized so they can be compared straightforwardly. By doing this we guarantee the accuracy of the data entered and readiness for the application of the time series analysis techniques. The first Early Data Analysis (EDA) step is to look for regularities, patterns, and possible anomalies. A visualization of these tools helps to see the time factors that are affecting the supply and demand of housing stocks.

Time series decomposition is employed to separate the data into its components: track, trend, seasonal, and remain. Therefore, it makes it easier to study long-term changes and periodic variations. The model should match the attributes determined in EDA, in this case. This can include traditional statistical approaches such as ARIMA or the more advanced machine learning ones like Long Short-Term Memory (LSTM) Networks. For particular vectors, estimations for parameters are carried out by historical data. This is done by fitting the model to the training set and by applying a set of parameters that result in the highest model's accuracy.

Economic variables, including GDP growth, interest rate, and unemployment rate, are combined in the model to react to outside economic effects. Demographic trends, such as a population surge and a shift in age ranges, are expected to reflect society's changing housing interests. The performance of the model is checked, if it is accurate using the holdout dataset, which was not used during the training phase. Measures like MAE (Mean Absolute Error) and RMSE (Root Mean Squared Error) are used for precision and reliability assessment.

Models' robustness is also assessed by performing sensitivity analysis that includes variation of significant variables. It makes sure that the model can learn in various scenarios and uncover warnings about risks and dangers. Results are reviewed against the Ankara housing market, which would give insights in main facts, determinants of fluctuations, possible future scenarios. The whole research did meticulously looks into the details and this helps to promote transparency and reproducibility as well.

The methodology will place the focus on the comprehensive overview of housing needs in Ankara, Turkey, which is achieved via a combined usage of the classical techniques of time series analysis and cutting-edge machine learning models. It plans on giving correct predictions by factoring in the multidimensional features of the local realty market.

3.1 Mathematical Model

With the exception of the special case combination of autoregressive (AR), integrated (I) , and moving average (MA), known as the ARIMA(p, d, q), namely the practical time series forecasting maths model is considered. Such a model is broadly employed in the areas of time series analysis and forecasting.

Supposing we label Y_t as the housing demand at time t.

Mathematical Model: Autoregressive Integrated Moving Average (ARIMA).

$$\left(\mathbf{1}-\boldsymbol{\theta_1 B}-\boldsymbol{\theta_2 B^2}-\cdots-\boldsymbol{\theta_p B^p}\right)(\mathbf{1}-\boldsymbol{B})^{\boldsymbol{d}}\boldsymbol{Y_t}\text{-}\boldsymbol{\epsilon_t}+\left(\mathbf{1}+\boldsymbol{\vartheta_1 B}+\boldsymbol{\vartheta_2 B^2}-\cdots+\boldsymbol{\vartheta_p B^p}\right).\boldsymbol{\epsilon_t} \qquad (1)$$

Here:

- $\boldsymbol{B}$ Operator of the back shift $(\boldsymbol{BY_t}=\boldsymbol{Y_{t-1}})$ shows the lag effect.
- d is the differencing degree to stationarity.

- $\boldsymbol{\theta_1, \theta_2, \dots, \theta_p}$ are the autoregressive coefficients.
- $\boldsymbol{\vartheta_1, \vartheta_2, \dots, \vartheta_p}$ are the moving average coefficients.
- . $\boldsymbol{\epsilon_t}$ is the white noise error term.

Model Components
1. Autoregressive (AR) Component:
$$\left(\mathbf{1} - \boldsymbol{\theta_1 B} - \boldsymbol{\theta_2 B^2} - \cdots - \boldsymbol{\theta_p B^p}\right) \tag{2}$$

The market model has been built in this way reflects the relationship between the current value and its previous values.

2. Integrated (I) Component:
$$(\mathbf{1} - \boldsymbol{B})^{\boldsymbol{d}} \tag{3}$$

Wherein stationarity is achieved through the means of differencing.

3. Moving Average (MA) Component:
$$\left(\mathbf{1} + \boldsymbol{\vartheta_1 B} + \boldsymbol{\vartheta_2 B^2} - \cdots + \boldsymbol{\vartheta_p B^p}\right) \tag{4}$$

The way, it works might be represented as a function of a past white noise error on the current value. Compute values $\boldsymbol{\theta_1, \theta_2, \dots, \theta_p}$, $\boldsymbol{\vartheta_1, \vartheta_2, \dots, \vartheta_p}$ by methods like maximum likelihood estimation (MLE). Test the model by calculating its errors (e.g. Mean absolute error, root mean squared error) and proof the residuals for randomness. The estimated ARIMA (p, d, q) model can be used to forecast future values of the housing demand series. This ARIMA model has the capability of having a flexible framework of time series forecasting and it can be modified if the behavior of housing demand data in Ankara, Turkey, is different somehow. Changes and upgrades can be constantly occurring, which might mean using different models, advanced methods, or machine learning for more precise predictions.

4 Case Study

Ankara, the capital of Turkey which has been undergoing these issues of population growth, urbanisation, and economic transformation has experienced dynamic changes in its real estate sector. This case study aims to carry out a time series analysis exhaustively so as to prepare a forecasting model somehow to predict housing demand in Ankara. Acquire housing sale figures in Ankara covering a period that shows property price fluctuations. Use variables like title where the property was sold, price trend changes, economic indicators, demographic factors, urban development data, and government policies as your basis to analyze the behavior of the real estate market. Do the data cleaning and preprocessing thoroughly. Taking care of missing values, and outliers, and making sure data consistency. With a datetime index, organize the dataset and standardize your data for analysis.

Conduct EDA to identify patterns, time limits, and possible outliers in the HOUSING demand data. Apply visualization tools to identify the temporal patterns of the market to better understand it. Write an essay about the role of social media influencers in the fashion industry. Discuss the impact of these influencers on consumer behavior, the challenges they face, and their long-term influence on the fashion industry. Use time series decomposition to decompose the data into a trend, season, and odd elements. This step thus helps to decode long-term tendencies and cyclical movements that affect, positive or negative, demand in housing. Based on EDA observation do it through a time series model selection. One option is to perform ARIMA, SARIMA, or any other sophisticated technique that could be suitable depending on the data properties. Work to identify optimal estimates for model parameters (p, d, q) using lag plots and time series plots involving double and partial autocorrelation functions. Employ the fitted time series model through a library such as Statsmodels. Fit model to training data and assess its performance using hold-out samples. Sometimes employ learning curve as validation. Ascertain that the model is reflecting what is, as a whole, predictable from the housing demand time series. Incorporate some economic indicators (such as GDP growth, interest rates, and unemployment rates) and additional demographic parameters (for example, population growth, and age distribution) into the model, so that it improves its predictability. Tailor the model so as to allow external factors to be included that change the disparity in housing demand. To stabilize the forecasting model using statistical quantitative tools such as Mean Absolute error (MAE) and Root Mean Squared error (RMSE). Residual series are analyzed for the randomness to spot if any signs of a well-fitted model are there or not. Apply the model, which is verified, to forecast future demand for housing in Ankara for the next few years. Delineate horizons and set fort a timeline with objectives of the analysis and tense of users in mind.

Explain the implications generated by the housing market Ankara forecasting outcomes. Tendency to mention prominent trends and discuss the major influencing factors that affect the housing demand and the projected trajectory. Make use of the results to offer advice to developers, investors, and policymakers regarding real estate investment.

Moreover, formulate the spaces for research and devising plans for further model enhancements. The given case study is a part of a scaffolded process of a time series analysis of housing demand in Ankara, Turkey. Its focus is the fact that data processing, use of the appropriate technique integration of relevant data, and ramifications of the forecasting process, are necessary to ensure the accuracy of the predictions.

The current study has taken a broader look at housing needs in Ankara, Turkey, and time series analysis has been utilized and we developed a model for the forecasts. The investigation was mostly geared towards supplying informed, crucial data for real estate players, investors, and makers who wanted to discover alternative ways of navigating through the dynamic real estate market in the capital city. In this work, we are trying to follow and calculate the numbers of house sales in Ankara province of the Turkey Republic over 4 years and applying the demand forecasting method to look into the future and forecast the number of sales in the following year. As a result, we choose specific ways for comparison and conscientiously select the most suitable technique. During carrying out the research work, the most adequate one for the given data will be chosen that is with time series methods, one of the forecasting techniques for the demand.

Using the Minitab program and with the help of Excel, Exponential Smoothing, Moving Average, Holt - Winters Methods were applied to the data. The results were then compared with each other. Their performances were analyzed. The reason for doing this is to find the most optimal solution and to make the demand forecast way most accurately. For the analysis part, Simple Linear Regression Analysis was used as a statistical method. For the solution of the problem mentioned above, the following steps were followed to reach the solution (Table 1).

Table 1. Solution Stages

Solution Stages	Work Performed	Solution Tool, Method, Technique, Program, etc.
1	Examining the existing system	Literature review, TurkStat data
2	Identification of problems or areas for improvement	Examination of data
3	Selection of the appropriate solution method	Literature review
4	Solving the problem	Minitab, Excel
5	Analysis	Multiple Regression Analysis
6	Comparison of results	Correlation analysis, Comparison of predictions

When the number of house sales in Ankara is analyzed from the available TURKSTAT data, 11744 houses were sold on average between 2014-2018, [1]. When we set aside the average of each month, January, February, June, and July stand out as being lower than the average. Except March, the highest number of house sales were made in the fall months (Table 2). The highest monthly average appears to be December. Data taken from data.tuik.gov.tr

Table 2. Monthly averages

Average Monthly Number of House Sales in Ankara Province between 2014-2018	
January	9825,8
February	10420,6
March	12760,2
April	11761,8
May	11955,6
June	11447,8
July	10578,2
August	11839,2
September	11860,4
October	12317,2
November	12079,8
December	14284,2
Overall Average	11744,23

Demand forecasting will be created using existing data rather than an area that needs to be developed. A literature review was conducted to select a solution method. Various foreign internet sources, online library databases, studies, articles, and theses on the subject were analyzed. As a result, Simple Exponential Smoothing, Moving Average, Holt, and Winter methods from Time Series methods were found to be the most suitable for the data. Currently, this is the housing sales data in Ankara between 2014-2018 (Figure 1). When we look at it, sales generally climb before January. This can be interpreted as people thinking that prices may increase after the New Year.

Fig. 1: Graph of Data between 2014-2018

The number of house sales in Ankara province between 2014 and 2018 obtained from TurkStat is taken into consideration and forecasts are generated with 3-month and 5-month moving averages, [1]. The Minitab program was used to find the results. With the 3-month Moving Average, there are not many close forecasts. When we look at the data, the Average Error Percentage is 17 and the Average Absolute Deviation is 1658 (Table 3).

Table 3. Month Moving Average Forecasts for Housing Sales

Month - Year	Realized Sale	Movable Average	Forecast	Error
Jan.19	6785	8856,67	8904,67	-2119,67
Feb.19	7690	8364,67	8856,67	-1166,67
Mar.19	10619	9061,67	8364,67	2254,33
Apr.19	8876	9636,33	9061,67	-185,67
May.19	9414	7906,67	9636,33	-222,33
Jun.19	5430	8111,67	7906,67	-2476,67
July.19	9491	8542,67	8111,67	1379,33
Aug.19	10707	11701,33	8542,67	2164,33
Sept.19	14906	13197,00	11701,33	3204,67
Oct.19	13978	14463,33	13197,00	781,00
Nov.19	14506	16189,33	14463,33	42,67
Dec.19	20084	16189,33	16189,33	3894,67
Measures of Accuracy				
MAPE			17	
MAD			1658	
MSD			4149241	

Table 4. Month Moving Average Forecasts for Housing Sales

Month - Year	Realized Sale	Movable Average	Forecast	Error
Jan.19	6785	9004,6	9566,8	-2781,8
Feb.19	7690	9213,0	9004,6	-1314,6
Mar.19	10619	8676,8	9213,0	1406,0
Apr.19	8876	8405,8	8676,8	199,2
May.19	9414	8766,0	8405,8	1008,2
Jun.19	5430	8783,6	8766,0	-3336,0
July.19	9491	9989,6	8783,6	707,4
Aug.19	10707	10902,4	9989,6	717,4
Sept.19	14906	12717,6	10902,4	4003,6
Oct.19	13978	14836,2	12717,6	1260,4
Nov.19	14506	14836,2	14836,2	-330,2
Dec.19	20084	14836,2	14836,2	5247,8
Measures of Accuracy				
MAPE		19		
MAD		1859		
MSD		5825783		

Forecasting with the 5-month Average has slightly more error than forecasting with the 3-month Average. It is confirmed by the fact that the Average Error Percentage is 19 and the Average Absolute Deviation is 1859 (Table 4 and Table 5).

Table 5. Estimation with α: 0.404686 for the Number of Housing Sales

Month - Year	Realized Sales	Exponential Adjustment	Forecast	Error
Jan.19	6785	9313,8394	11032,9084	-4247,9084
Feb.19	7690	8656,6943	9313,8394	-1623,8394
Mar.19	10619	9450,8119	8656,6943	1962,3057
Apr.19	8876	9218,1936	9450,8119	-574,8119
May.19	9414	9297,4337	9218,1936	195,8064
Jun.19	5430	7732,3374	9297,4337	-3867,4337
July.19	9491	8444,0435	7732,3374	1758,6626
Aug.19	10707	9359,8303	8444,0435	2262,9565
Sept.19	14906	11604,2876	9359,8303	5546,1697
Oct.19	13978	12564,8958	11604,2876	2373,7124
Nov.19	14506	13350,4335	12564,8958	1941,1042
Dec.19	20084	16075,4136	13350,4335	6733,5665
Measures of Accuracy				
MAPE			27	
MAD			2757	
MSD			14425965	

In this section, the exponential smoothing method is used. At first, the smoothing coefficient was left to the Minitab program. The program automatically took the value as α : 0.404686. Then, α manually entered as 0.2, 0.5, 0.7, and other estimates were added as Table 6, Table 7 and Table 8.

Table 6. Estimation for the Number of Housing Sales with α : 0.2

Month - Year	Realized Sales	Exponential Adjustment	Forecast	Error
Jan.19	6785	10279,8108	11153,5135	-4368,5135
Feb.19	7690	9761,8487	10279,8108	-2589,8108
Mar.19	10619	9933,2789	9761,8487	857,1513
Apr.19	8876	9721,8231	9933,2789	-1057,2789
May.19	9414	9660,2585	9721,8231	-307,8231
Jun.19	5430	8814,2068	9660,2585	-4230,2585
July.19	9491	8949,5654	8814,2068	676,7932
Aug.19	10707	9301,0524	8949,5654	1757,4346
Sept.19	14906	10422,0419	9301,0524	5604,9476
Oct.19	13978	11133,2335	10422,0419	3555,9581
Nov.19	14506	11807,7868	11133,2335	3372,7665
Dec.19	20084	13463,0294	11807,7868	8276,2132
Measures of Accuracy				
MAPE			29	
MAD			3055	
MSD			14425965	

Although the four estimates are not very close to each other, their margins of error are similar. The last estimate, with an Average Percentage Error of 17 and an Average Absolute Deviation of 1814, and a smoothing coefficient α : 0.7, seems to be the most distant from the actual data.

Table 7. Estimation for the Number of Housing Sales with α : 0.5

Month - Year	Realized Sales	Exponential Adjustment	Forecast	Error
Jan.19	6785	8921,2014	11057,4028	-4272,4028
Feb.19	7690	8305,6007	8921,2014	-1231,2014
Mar.19	10619	9462,3004	8305,6007	2313,3993
Apr.19	8876	9169,1502	9462,3004	-586,3004
May.19	9414	9291,5751	9169,1502	244,8498
Jun.19	5430	7360,7875	9291,5751	-3861,5751
July.19	9491	8425,8938	7360,7875	2130,2125
Aug.19	10707	9566,4469	8425,8938	2281,1062
Sept.19	14906	12236,2234	9566,4469	5339,5531
Oct.19	13978	13107,1117	12236,2234	1741,7766
Nov.19	14506	13806,5559	13107,1117	1398,8883
Dec.19	20084	16945,2779	13806,5559	6277,4441
Measures of Accuracy				
MAPE			26	
MAD			2640	
MSD			10257135	

Table 8. Estimation for the Number of Housing Sales with α : 0.7

Month - Year	Realized Sales	Exponential Adjustment	Forecast	Error
Jan.19	6785	8125,8672	11254,5574	-4469,5574
Feb.19	7690	7820,7602	8125,8672	-435,8672
Mar.19	10619	9779,5281	7820,7602	2798,2398
Apr.19	8876	9147,0584	9779,5281	-903,5281
May.19	9414	9333,9175	9147,0584	266,9416
Jun.19	5430	6601,1753	9333,9175	-3903,9175
July.19	9491	8624,0526	6601,1753	2889,8247
Aug.19	10707	10082,1158	8624,0526	2082,9474
Sept.19	14906	13458,8347	10082,1158	4823,8842
Oct.19	13978	13822,2504	13458,8347	519,1653
Nov.19	14506	14300,8751	13822,2504	683,7496
Dec.19	20084	18349,0625	14300,8751	5783,1249
Measures of Accuracy				
MAPE			25	
MAD			2463	
MSD			9522205	

4.1 Holt-Winters Method

In Holt-Winters exponential smoothing methods, trend and seasonality are in the foreground when forecasting. Each value in the series is estimated with a separate equation. There are two methods in the Holt-Winters method: additive and multiplicative. Both methods are based on three equations. The first one is used to determine the level of the series in period t, the second one is used to determine the trend of the series and the third one is used to determine the seasonal component (Table 9 and Table 10).

Equations for the multiplicative method:

$$F_{t+m} = (L_t + b_t m) S_{t-s+m} \tag{5}$$

$$L_t = \alpha \left(\frac{Y_t}{S_{t-s}} \right) + (1-\alpha)(L_{t-1} + b_{t-1}) \tag{6}$$

$$b_t = \beta (L_t - L_{t-1}) + (1-\beta) b_{t-1} \tag{7}$$

$$S_t = \gamma \left(\frac{Y_t}{L_t} \right) + (1-\gamma) S_t \tag{8}$$

$\boldsymbol{L_t}$: General level of the series in period t
$\boldsymbol{Y_t}$: Observation value
$\boldsymbol{S_t}$: Seasonal component
$\boldsymbol{b_t}$: Trend component
$\boldsymbol{\alpha}$: Level correction constant
$\boldsymbol{\beta}$: Trend correction constant
$\boldsymbol{\gamma}$: Seasonal correction constant
$\boldsymbol{F_{t+m}}$: Forecast value for the forward period

Equations for the additive method:

$$L_t = \alpha (Y_t - S_{t-s}) + (1-\alpha)(L_{t-1} + b_{t-1}) \tag{9}$$

$$b_t = \beta (L_t - L_{t-1}) + (1-\beta) b_{t-1} \tag{10}$$

$$S_t = \gamma_t (Y_t - L_t) + (1-\gamma) S_{t-s} \tag{11}$$

$$F_{t+m} = L_t + b_t m + S_{t-s+m} \tag{12}$$

Table 9. Demand Forecasting with Holt-Winters Multiplicative Method

Month - Year	Realized Sales	Exponential Adjustment	Baseline (α)	Trend (β)	Season (γ)	Estimation	Error
Jan.19	6785	8246,7793	9182,9177	-90,3168	0,7969	8198,1968	-1413,1968
Feb.19	7690	7957,5308	8997,1030	-94,3308	0,8627	7879,2661	-189,2661
Mar.19	10619	9499,8281	9407,4672	-73,1173	1,0799	9400,2265	1218,7735
Apr.19	8876	8041,9173	9290,8412	-74,9461	0,9592	8971,6412	-95,6412
May.19	9414	9108,8232	9384,3053	-67,8674	0,9880	9036,3373	377,6627
Jun.19	5430	8534,9209	7853,4132	-129,3618	0,8376	8473,1962	-3043,1962
July.19	9491	6652,8989	9260,2587	-84,7922	0,9041	6523,6416	2967,3584
Aug.19	10707	8594,4565	10219,0177	-21,7692	0,9676	8534,1226	2172,8774
Sept.19	14906	10127,9867	12314,6788	67,2311	1,0634	10106,4114	4799,5886
Oct.19	13978	12815,7281	12840,8362	86,5208	1,0565	12885,6947	1092,3053
Nov.19	14506	12544,7321	13767,3112	121,8260	1,0022	12629,2578	1876,7422
Dec.19	20084	16793,7888	15015,2216	160,1579	1,2587	16942,3959	3141,6041
Measures of Accuracy							
MAPE				18			
MAD				1866			
MSD				5330828			

Table 10. Demand Forecasting with Holt-Winters Aggregation Method

Month - Year	Realized Sales	Exponential Adjustment	Baseline (α)	Trend (β)	Season (γ)	Estimation	Error
Jan.19	6785	7339,7181	8536,1883	-615,1393	-1784,1468	6716,3372	68,6628
Feb.19	7690	7234,8918	8349,2279	-486,0858	-1173,0427	6619,5526	1070,4474
Mar.19	10619	9315,7413	8578,5200	-271,8823	1183,3087	8829,0557	1789,9443
Apr.19	8876	8353,8647	8614,2388	-176,6090	-129,3720	8081,9724	794,0276
May.19	9414	8513,4325	8878,5004	-67,3478	18,4549	8336,8235	1077,1765
Jun.19	5430	7851,3989	7881,6129	-332,2097	-1312,1654	7803,8491	-2373,8491
July.19	9491	6159,5532	9014,8658	107,4291	-1282,4209	5827,3435	3663,6565
Aug.19	10707	8562,4718	9937,1348	351,8810	-207,9421	8669,9009	2037,0991
Sept.19	14906	10146,1240	12052,2135	880,8404	737,9488	10498,0050	4407,9950
Oct.19	13978	12829,8881	13036,9625	912,9130	809,7472	13710,7285	267,2715
Nov.19	14506	13558,8799	13966,5583	917,0178	523,0222	14471,7928	34,2072
Dec.19	20084	17190,0485	19674,3492	1154,2487	3460,7231	18107,0673	1976,9327
Measures of Accuracy							
MAPE				18			
MAD				1630			
MSD				4398088			

Looking at the two tables, we can say that both forecasting methods perform well. On the other hand, the Aggregate Method stands out here by making slightly more realistic forecasts. Especially since the sales figures for December 2019 are abnormal, other forecasting methods cannot come very close to predicting this sale.

4.2 Holt's Linear Method

It is generally used when there is a trend in the time series but no seasonality. This method again uses simple exponential smoothing. The following equation is used to determine the fundamental level of the trend:

$$L_t = \alpha Y_t + (1-\alpha)(L_{t-1} + T_{t-1}) \quad (13)$$

$$F_{t+n} = L_t + nT_t \quad (14)$$

$$T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1} \quad (15)$$

$\boldsymbol{L_t}$: Expected level in period t
$\boldsymbol{\alpha}$: Smoothing coefficient of the level
$\boldsymbol{Y_t}$: Realized value in the period
$\boldsymbol{T_t}$: trend in period t
$\boldsymbol{\beta}$: Smoothing coefficient of the trend
$\boldsymbol{n}$: Number of foreseeable periods

As a result of the data, Holt's Linear Method was used in the Minitab program (Table 11 and Table 12).

Table 11. Demand Forecasting with Holt's Linear Method

Month - Year	Realized Sales	Exponential Adjustment	Baseline (α)	Trend (β)	Season (γ)	Estimation
Jan.19	6785	7900,5336	7900,5336	-111,6979	11305,9554	-4520,9554
Feb.19	7690	7714,3874	7714,3874	-113,6825	7788,8357	-98,8357
Mar.19	10619	9874,2487	9874,2487	-53,6655	7600,7249	1018,2751
Apr.19	8876	9100,0734	9100,0734	-72,4418	9820,5832	-944,5832
May.19	9414	9320,8853	9320,8853	-64,9405	9036,6315	377,3685
Jun.19	5430	6374,0416	6374,0416	-140,9924	9255,9448	-3825,9448
July.19	9491	8687,1094	8687,1094	-76,2311	6233,0492	3257,9508
Aug.19	10707	10189,7876	10189,7876	-34,5645	8610,8783	2096,1217
Sept.19	14906	13733,7585	13733,7585	59,8710	10155,2231	4750,7769
Oct.19	13978	13932,5071	13932,5071	63,5359	13793,6298	184,3706
Nov.19	14506	14380,1695	14380,1695	73,6728	13996,0430	509,9570
Dec.19	20084	18694,7737	18694,7737	185,5886	14455,8423	5630,1577
Measures of Accuracy						
MAPE			25			
MAD			2435			
MSD			9568822			

Although Holt's Linear Method gives very close forecasts for some months, we can easily state that it is not one of the most appropriate forecasting methods for the data we have, as it gives values that are very far from the truth in general. In the regression method, observations of the dependent and independent variables are needed to find the coefficients to be used. Eq. 16 represent s the main mass regression method and eq. 17 shows the sample regression model). A linear regression equation between two variables can be expressed.

$$Y = \beta_0 + \beta_1 X + \varepsilon \quad (16)$$

$$= \alpha + \mathrm{b}X + e_i \quad (17)$$

$\boldsymbol{Y}$: Dependent variable
$\boldsymbol{X}$: Independent variable
Ý: Estimated value
$\boldsymbol{\beta_0}$ ***ve*** $\boldsymbol{\beta_1}$: Parameters of the regression equation
$\boldsymbol{\alpha}$ ***ve b***: Coefficients of the estimated regression equation (estimators of β_0 and β_1)
$\boldsymbol{\varepsilon}$: Error term, e: Error estimator

Table 12. Demand Forecasting with Linear Regression Method

Month - Year	Realized Sales	Exponential Adjustment	Forecast
Jan.19	6785	9043,41	-2258,4
Feb.19	7690	9962,93	-2272,9
Mar.19	10619	13002,4	-2383,4
Apr.19	8876	11314	-2438,0
May.19	9414	11716,2	-2302,2
Jun.19	5430	9612,22	-4182,2
July.19	9491	10542,1	-1051,1
Aug.19	10707	11430,6	-723,6
Sept.19	14906	12771,6	2134,4
Oct.19	13978	12674,3	1303,7
Nov.19	14506	11915,7	2590,3
Dec.19	20084	15502,9	4581,1
Measures of Accuracy			
MAPE		24	
MAD		2352	
MSD		6688499	

The forecasts made with the Linear Regression Method do not deviate too far from the actual data, but they do not get too close to the actual data either. The Average Error Percentage is 24, which seems to be the highest value in the forecasts so far. The prediction was made with the NeuroXL Predictor add-in via Microsoft Excel. Accuracy measures were then found according to the results (Table 13).

Table 13. Demand Forecasting with Artificial Neural Networks

Month - Year	Realized Sales	Exponential Adjustment	Forecast
Oca.19	6785	11158	-4372,9
Şub.19	7690	11194	-3503,7
Mar.19	10619	11460	-841,4
Nis.19	8876	11276	-2399,6
May.19	9414	11326	-1912,3
Haz.19	5430	11135	-5705,3
Tem.19	9491	11334	-1843,2
Ağu.19	10707	11471	-763,6
Eyl.19	14906	12278	2627,6
Eki.19	13978	11814	2164,3
Kas.19	14506	11983	2522,9
Ara.19	20084	16897	3187,4
Measures of Accuracy			
MAPE		30	
MAD		2654	
MSD		8847166	

Artificial Neural Networks generally predicted the same values. The exception here is the relatively close prediction of December.

5 Analysis

Demand forecasting applications with actual sales amounts from the Minitab program were analyzed by multiple regression analysis methods. Since more than 5 analyses could not be performed in multiple regression analysis with the Minitab program, the analyses were performed piece by piece. Multiple Regression Analyses between 3-Month and 5-Month Average and Actual Values are considered. Among the two, the 3-month moving average method is closer to the actual data. Considering that the R-sq value is 81.21%, we can also understand the conformity of the forecast values to the actual sales values (Figure 2).

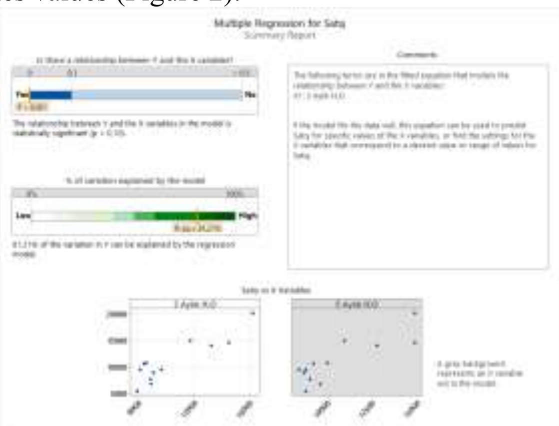

Fig. 2: Multiple Regression Analysis with 3 and 5 Month Moving Average

In the analysis, $p < 0.001$ indicates that it is statistically significant (Figure 3).

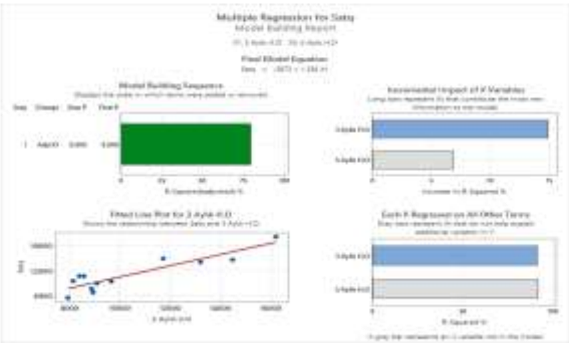

Fig. 3: Multiple Regression Analysis with 3 and 5-Month Moving Average Continued

Multiple Regression Analysis between Exponential Adjustment Methods and Actual Values

There are 4 separate tables of prediction values for Exponential Adjustment Methods. Since the actual values cannot be analyzed with all four in the Minitab program, they are divided into two (Figure 4 and Figure 5).

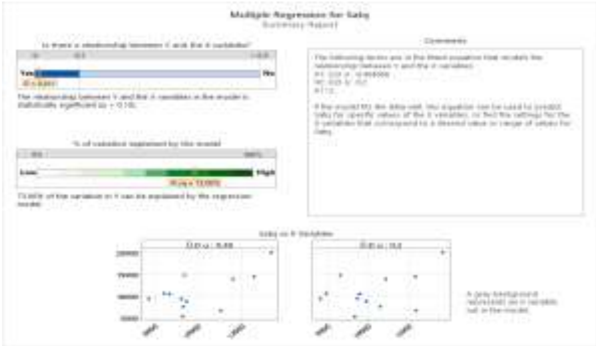

Fig. 4: Exponential Flat. Multiple Regression Analysis for α : 0.404686 and α : 0.2

0.4046 is the smoothing value given by the minitab program as mentioned before. R-sq was found to be 73.06%. It is lower than the previous method.
p = 0.011. It is seen that it is below the threshold value of $p < 0.10$.

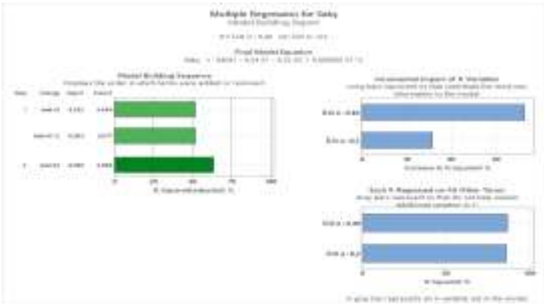

Fig. 5: Exponential Flat. Multiple Regression Analysis for α : 0.404686 and α : 0.2

Starting from Figure 6, α : 0.5 and α : 0.7 Multiple Regression Analyses are shown.

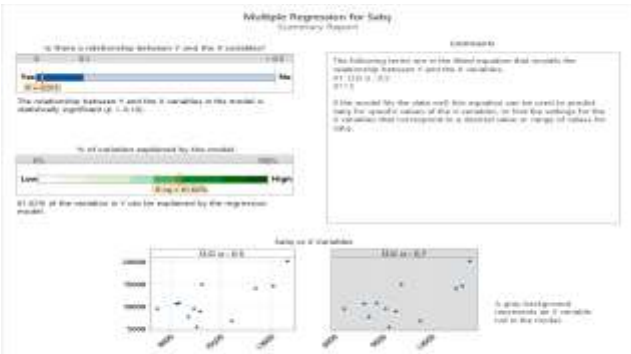

Fig. 6: Exponential Flat. α : 0.5 and α : 0.7 Multiple Regression Analysis

Here, even though the α : 0.5 estimate is closer to the actual values and the p-value is 0.013, R-sq = 61.62%, which means that its conformity to the actual data is lower (Figure 7).

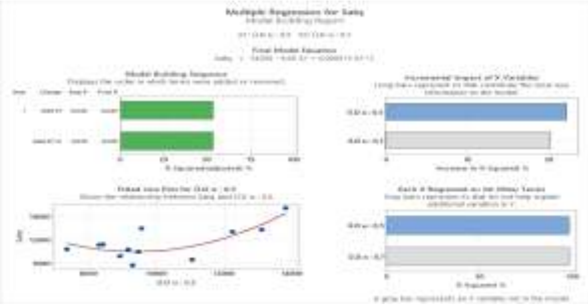

Fig. 7: Exponential Flat. α : 0.5 and α : 0.7 Multiple Regression Analysis

Holt-Winters methods have two separate estimation tables, Multiplicative and Additive. The actual values and the multiple regression analysis between them are given together (Figure 8).

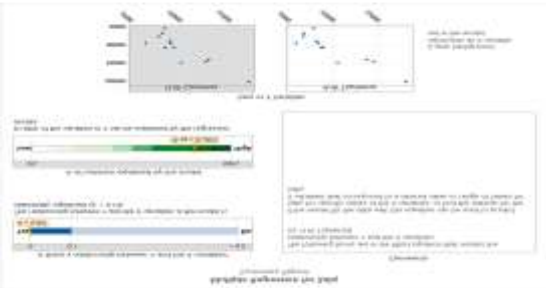

Fig. 8: Holt-Winters Multiplicative and Additive Multiple Regression Analysis

Here, the Holt-Winters Additive Method stands out as being more appropriate to the actual data than the other method. R-sq = 81.68% confirms this. $p < 0.001$ indicates that it is statistically significant (Figure 9).

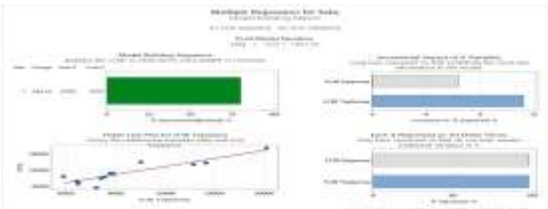

Fig. 9: Holt-Winters Multiplicative and Additive Multiple Regression Analysis

Figure 9 shows how close the Holt-Winters Summative Method is to the actual values, as can be seen from the bottom left graph. Multiple regression analysis was applied to the predicted values and actual values made with these three methods together. As can be seen in Figure 10, R-sq = 89.31. $p < 0.001$ indicates that the result is statistically significant (Figure 10 and Figure 11).

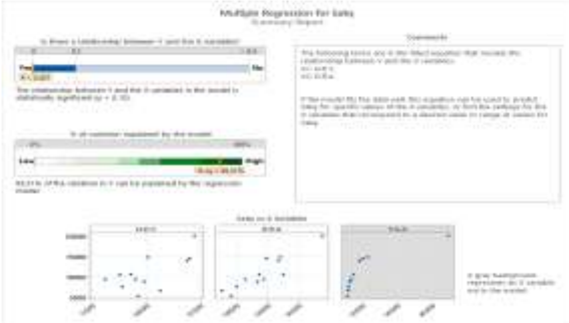

Fig. 10: Multiple Regression Analysis with Three Methods

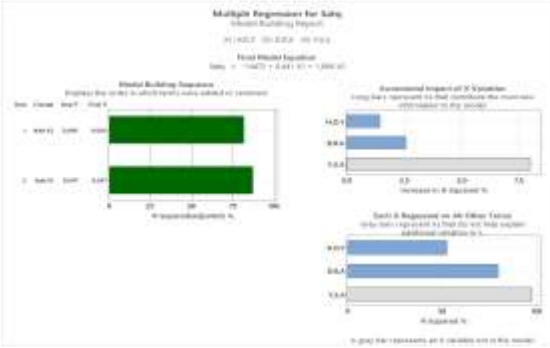

Fig. 11: Multiple Regression Analysis with Three Methods

From an economic point of view, a close estimate of the number of house sales is not a direct indicator of whether house prices will rise or fall, but it can be useful for real estate investors to predict market trends for people considering buying a house. These forecasts, no matter how close to reality they are, should be taken into account as investment advice. When evaluating the results, it is first necessary to evaluate the accuracy measures of all our forecasting methods in a table. MAPE - Mean Absolute Deviation, MAD - Mean Absolute Error Percentage, and MSD - Mean Square Error values of the estimation methods used in the study were compared (Table 14).

Table 14. Comparison of Accuracy Values of Forecasting Methods

	MAPE	MAD	MSD
3 Month Moving Average	17	1658	4149241
5 Month Moving Average	19	1859	5825783
Exponential Smoothing α : 0.4046	27	2757	11131266
Exponential Correction α : 0.2	29	3055	14425965
Exponential Correction α : 0.5	26	2640	10257135
Exponential Correction α : 0.7	25	2463	9522205
Holt-Winters Multiplicative	18	1866	5339828
Holt-Winters Additive	**16**	**1630**	**4396988**
Holt's Naturalistic Method	25	2435	9566822
Linear Regression Analysis	24	2352	6688499
Artificial Neural Networks	30	2654	8847166

When we compare the accuracy values, Holt-Winters Aggregation Method stands out with its reliability in both MAPE and MAD values. The closest to it are the forecast values made with the 3-month Moving Average.

Finally, the correlation analysis of actual sales values and forecast values was performed. Since using all of them together in the Minitab program was a bit complicated, the process was carried out in groups of 2 and 3 (Table 15).

Table 15. Correlation of All Demand Forecasts with Sales Values

	Correlation
3 Month M.A.	0,901
5 Month M.A.	0,854
Ü.D α : 0,404686	0,638
Ü.D α : 0,2	0,424
Ü.D α : 0,5	0,66
Ü.D α : 0,7	0,665
H.W Multiplicative	0,876
H.W Total	0,904
H.D.Y	0,662
D.R.A	0,861
Y.S.A	0,824

The Holt-Winters Additive Method was found to be the closest to the 1. 3-Month Moving Average and the Holt-Winters Multiplicative Method were also found to be the next closest values (Figure 12 and Figure 13).

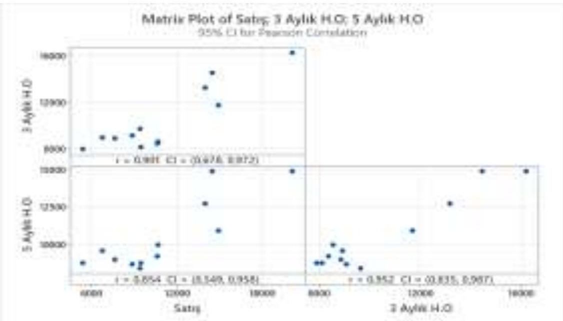

Fig. 12: 3 and 5-Month Moving Average Correlation Analysis

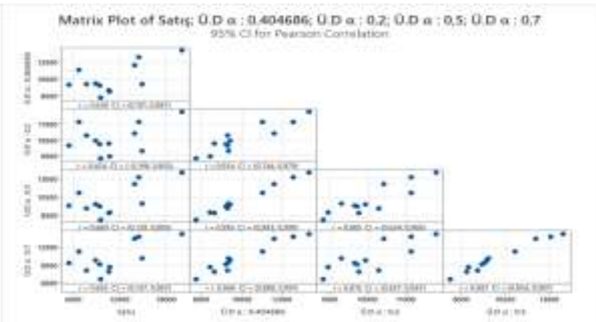
Fig. 13: Exponential Correlation Analysis

As a result, the time series analysis revealed intricate patterns and seasonality within the housing demand data, emphasizing the need for a tailored forecasting approach. Integration of economic indicators and demographic factors significantly improved the predictive accuracy of the forecasting model.

The selected forecasting model, whether based on ARIMA, SARIMA, or other advanced techniques, demonstrated its ability to capture and forecast housing demand trends. Limitations of the study, while the forecasting model presented accurate predictions, it is essential to acknowledge the inherent uncertainties in real estate markets. External factors, such as global economic shifts or unforeseen events, may influence the accuracy of predictions. Future work could explore the integration of machine learning algorithms and explore more granular spatial analysis to provide localized forecasts for different districts within Ankara (Figure 14 and Figure 15).

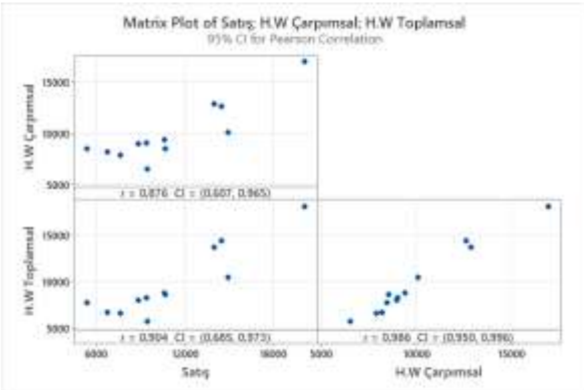

Fig. 14: Holt-Winters Method Correlation Analysis

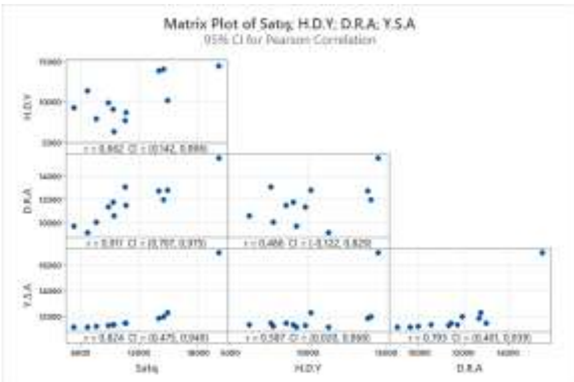

Fig. 15: Holt's D. Y. - D. Reg. - Artificial Neural Networks Correlation Analysis

6 Conclusion

This study contributes to the body of knowledge in real estate forecasting by combining time series analysis with economic and demographic factors in the specific context of Ankara, Turkey. The methodology and results form a valuable source for researchers and practitioners trying to forecast housing demand, particularly in regions with housing demands driven by rapid urbanization. The main apprentices of the developed forecasting model have a great opportunity to improve the level of their strategic decisions, investment optimizations, and real estate market developments that line up with the changing requests of Ankara. The visible description of the process that has been used means that the results of the study can be repeated by others which enables the reproduction and application of similar methods in other locations.

As a result, this study has set a solid step for determining and forecasting residential demand in Ankara, Turkey. The success of time series in combination with economic and demographic issues suggests the application of an integrated model that captures different dimensions of the real estate market. Along with the process of Ankara's urban transformation, the results of our research will suggest future courses of action for stakeholders to make their decisions on a solid basis within the always-changing real estate market.

References:

[1] House Sales Statistics, May 2022. TURKSTAT, [Online]. https://data.tuik.gov.tr/Bulten/Index?p=House-Sales-Statistics-May-2022-45677&dil=2 (Accessed Date: March 3, 2024).

[2] Huang, M.C., Time-varying diversification strategies: The roles of state-level housing assets in optimal portfolios, *International Review of Economics & Finance*, Vol. 55, May 2018, pp.145-172.

[3] Kim, H., Park, S., Kim, S., Time-series clustering and forecasting household electricity demand using smart meter data, *Energy Reports*, Vol. 9, December 2023, pp.4111-4121

[4] Banas, J., Utnik-Banas, K., Evaluating a seasonal autoregressive moving average model with an exogenous variable for short-term timber price forecasting, *Forest Policy and Economics*, Vol. 131, October 2021, 102564

[5] Yii, K.J., Tan, C.T., Ho, W.K., Kwan, X.H., Nerissa, F.S., Tan, Y.Y., Wong, K.H., Land availability and housing price in China: Empirical evidence from nonlinear autoregressive distributed lag (NARDL), *Land Use Policy*, Vol. 113, February 2022, 105888

[6] Lu,Y., Li, J., Yang, H., Time-varying impacts of monetary policy uncertainty on China's housing market, *Economic Modelling*, Vol. 118, January 2023, 106081

[7] Cipollini, A., Paria, F., Housing market shocks in italy: A GVAR approach, *Journal of Housing Economics*, Vol. 50, December 2020, 101707

[8] Biswas, S., Chakrabortty, R.K., Turan, H.H., Elsawah, S., Consideration of uncertainties in a dynamic modeling system integrated with a deep learning based forecasting approach, *CIRP Journal of Manufacturing Science and Technology*, Vol. 44, September 2023, Pages 27-44

[9] Kim, S., Innovating knowledge and information for a firm-level automobile demand forecast system: A machine learning perspective, *Journal of Innovation & Knowledge*, Vol. 8, Issue 2, April–June 2023, 100355

[10] Javanmard, M.E., Ghaderi, S.F., Energy demand forecasting in seven sectors by an optimization model based on machine learning algorithms, *Sustainable Cities and Society*, Vol. 95, August 2023, 104623

[11] Zhan, C., Liu, Y., Wu, Z., Zhao, M., Chow, T.W.S., A hybrid machine learning framework for forecasting house price, *Expert Systems with Applications*, Vol. 233, 15 December 2023, 120981

[12] Bhavsar, S., Pitchumani, R., Reynolds, M., Merket, N., Reyna, J., Machine learning surrogate of physics-based building-stock simulator for end-use load forecasting, *Energy and Buildings*, Vol. 296, 1 October 2023, 113395

[13] Chang, L., Mohsin, M., Hasnaoui, A., Taghizadeh-Hesary, F., Exploring carbon dioxide emissions forecasting in China: A policy-oriented perspective using projection pursuit regression and machine learning models, *Technological Forecasting and Social Change*, Vol. 197, December 2023, 122872.

[14] Stavinova, E., Timoshina, A., Chunaev, P., Forecasting the volume of mortgage loans with open Internet data in the period of noticeable changes in the Russian mortgage market, *Procedia Computer Science*, Vol. 193, 2021, pp.266-275

[15] Suh, H., Regionally heterogeneous housing cycles and housing market stabilization

policies: Evidence from Korea, *Economic Modelling*, Vol. 120, March 2023, 106192.

[16] Zhao, S., Huo, T., Chen, L., The influence of multi-level factors on Chinese residents' purchase decisions of green housing: A system dynamics approach, *Sustainable Cities and Society*, Vol. 99, December 2023, 105001.

[17] Ali, C.H., Roy, D., Amireche, L., Antoni, J.P., Development of a Cellular Automata-based model approach for sustainable planning of affordable housing projects: an application case study in Algiers, *Land Use Policy*, Vol. 125, February 2023, 106468.

[18] Ge, X., Li, X.L., Li, Y., Liu, Y., The driving forces of China's business cycles: Evidence from an estimated DSGE model with housing and banking, *China Economic Review*, Vol. 72, April 2022, 101753.

[19] Hui, E.C., Yu, K., Ng, I.M., The dynamics of housing demand under a linked-exchange rate system, *Habitat International*, Vol. 44, October 2014, Pages 50-61.

[20] Ahmed, A.I., Mcleod, R.S., Gustin, M., Forecasting underheating in dwellings to detect excess winter mortality risks using time series models, *Applied Energy*, Vol. 286, 15 March 2021, 116517.

[21] Rapach, D.E., Strauss, J.K., Differences in housing price forecastability across US states, *International Journal of Forecasting*, Vol. 25, Issue 2, April–June 2009, Pages 351-372.

[22] Cheng, Q., Argon, N.T., Evans, C.S., Liu, Y., Platts-Mills, T.F., Ziya, S., Forecasting emergency department hourly occupancy using time series analysis, *The American Journal of Emergency Medicine*, Vol. 48, October 2021, Pages 177-182.

[23] Han, Y., Tian, Y., Yu, L., Gao, Y., Economic system forecasting based on temporal fusion transformers: Multi-dimensional evaluation and cross-model comparative analysis, *Neurocomputing*, Vol. 552, 1 October 2023, 126500.

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

- C.Kanbolat, S.Turgay, E.S. Kara – investigation,
- C.Kanbolat, E.S.Kara - validation and
- S.Turgay writing & editing.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

No funding was received for conducting this study.

Conflict of Interest

The authors have no conflicts of interest to declare.