# DEVELOPING A GIS TOOL TO ANALYZE HOUSING PRICE VARIABILITY IN URBAN REGIONS CASE STUDY: ANKARA

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

DUAA ABU SADAA

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
GEODETIC AND GEOGRAPHIC INFORMATION TECHNOLOGIES

NOVEMBER 2024

Approval of the thesis:

**DEVELOPING A GIS TOOL TO ANALYZE HOUSING PRICE VARIABILITY IN URBAN REGIONS CASE STUDY: ANKARA**

Submitted by **DUAA ABU SADAA** in partial fulfillment of the requirements for the degree of **Master of Science** i**n Geodetic and Geographic Informatıon Technologies, Middle East Technical University** by,

Prof. Dr. Naci Emre ALTUN
Dean, Graduate School of **Natural and Applied Sciences** _______________

Prof.Dr. Zuhal Akyürek
Head of the Department, **GGIT** _______________

Prof.Dr. Zuhal Akyürek
Supervisor, **GGIT, METU** _______________

Dr. Tuncay Küçükpehlivan
Co-Supervisor, **BAŞARSOFT** _______________

**Examining Committee Members:**

Prof. Dr. Hediye Tüydeş
Civil Eng, METU _______________

Prof. Dr. Zuhal Akyürek
Civil Eng, METU _______________

Prof. Dr. Müge Akkar Ercan
City and Regional Planning, METU _______________

Prof. Dr. Ebru Vesile Öcalir
City and Regional Planning, Gazi Univ. _______________

Prof.Dr. Aslı Özdarıcı OK
Land Registry and Cadastre., HaciBayramVeli U _______________

Date: 28.11.2024

**I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.**

Name Last name: Duaa Abu Sadaa

Signature:

## ABSTRACT


**DEVELOPING A GIS TOOL TO ANALYZE HOUSING PRICE VARIABILITY IN URBAN REGIONS CASE STUDY: ANKARA**

Abu Sadaa, Duaa
Master of Science, Geodetic and Geographic Information Technologies
Supervisor: Prof. Dr. Zuhal Akyürek
Co-Supervisor: Dr. Tuncay Küçükpehlivan

November 2024, 165 pages

Ankara's population has grew by 5.35%, from 5,118,000 in 2020 to 5,397,000 in 2023, expanding the inhabited areas and requiring advanced urban analysis tools. This study introduces a customized Geographic Information System (GIS) application, developed with MapInfo, Visual Studio, and C#, to analyze and visualize housing price distributions through heat maps, Thiessen polygons, and comparison tools that reveal how various factors influence housing prices. The GIS application is a key tool for urban planners, real estate developers, and researchers, offering features like visualizing histograms of variables across price intervals and districts, and options to sort and zoom into specific neighborhoods. These tools enable a detailed exploration of spatial heterogeneity in the housing market, clarifying complex dynamics.

It is aimed to present the GIS application's analytical capabilities by integrating advanced spatial analysis methods, including Ordinary Least Squares (OLS) and Geographically Weighted Regression (GWR). These models reveal the relationships between housing prices and factors like proximity to transportation, hospitals, universities, shopping malls, banks, schools, and supermarkets. The results, visualized through the application, confirm spatial heterogeneity in housing prices across Ankara, particularly higher prices in the western regions due to key amenities such as proximity to transportation. The application highlights Çankaya’s unique characteristics, requiring separate analysis. As well as, a local analysis of Kızılay and Çayyolu shows that factors like transportation access and proximity to schools drive higher housing prices, especially in Çayyolu. This emphasizes the importance of neighborhood-level investigations to understand the dynamics influencing housing prices in different areas of Ankara.

Keywords: GIS, Spatial Analysis, Housing Prices, Spatial Heterogeneity, Türkiye.

# ÖZ

## KENTSEL BÖLGELERDE KONUT FİYAT DEĞİŞKENLİĞİNİ ANALİZ ETMEK İÇİN BİR GIS ARACI GELİŞTİRME VAKA ÇALIŞMASI: ANKARA

Abu Sadaa, Duaa
Yüksek Lisans, Jeodezi ve Coğrafi Bilgi Teknolojileri
Tez Yöneticisi: Prof. Dr. Zuhal Akyurek
Ortak Tez Yöneticisi: Doç. Dr. Tuncay Kucukpehlivan

Kasım 2024, 165 sayfa

Ankara'nın nüfusu 2020 yılında 5.118.000'den 2023 yılında 5.397.000'e, %5,35 oranında artmış, yerleşim alanlarını genişletmiş ve ileri düzey kentsel analiz araçlarına ihtiyaç duyulmuştur. Bu tez, konut fiyatı dağılımlarını analiz etmek ve görselleştirmek için MapInfo, Visual Studio ve C# kullanılarak geliştirilen özelleştirilmiş bir Coğrafi Bilgi Sistemi (CBS) uygulamasını tanıtmaktadır. Isı haritaları, Thiessen poligonları ve çeşitli faktörlerin konut fiyatlarını nasıl etkilediğini ortaya koyan karşılaştırma araçları ile analizler yapılmaktadır. CBS uygulaması, fiyat aralıkları ve ilçeler arasındaki değişkenlerin histogramlarını görselleştirme, belirli mahalleleri sıralama ve yakınlaştırma gibi özellikler sunarak şehir planlamacıları, gayrimenkul geliştiricileri ve araştırmacıları için önemli bir araç sunmaktadır. Bu araçlar, konut piyasasındaki mekânsal heterojenliği ayrıntılı bir şekilde inceleyerek karmaşık dinamiklerin daha iyi anlaşılmasını sağlamaktadır.

CBS uygulamasının analitik yeteneklerini gelişmiş mekânsal analiz yöntemleri olan En Küçük Kareler (OLS) ve Coğrafi Ağırlıklı Regresyon (GWR) modellezi entegre edilerek geliştirmek amaçlanmıştır. Bu modeller, konut fiyatları ile ulaşım, hastaneler, üniversiteler, alışveriş merkezleri, bankalar, okullar ve süpermarketler gibi faktörler arasındaki ilişkileri ortaya koymaktadır. Uygulama aracılığıyla

görselleştirilen sonuçlar, Ankara genelinde konut satış verilerinde mekânsal heterojenliği doğrulamakta, özellikle ulaşım gibi önemli olanaklara yakınlık nedeniyle batı bölgelerinde daha yüksek konut fiyatları görülmektedir. Uygulama, Çankaya ilçesinin kendine özgü özelliklerini vurgulayarak, ayrı bir analiz yapılmasını gerektirmiştir. Ayrıca, Kızılay ve Çayyolu'nun yerel analizleri, ulaşım erişimi ve okullara yakınlık gibi faktörlerin özellikle Çayyolu'nda konut fiyatlarını arttırdığını göstermektedir. Bu, Ankara'nın farklı bölgelerindeki konut fiyatlarını etkileyen dinamikleri anlamak için mahalle düzeyinde incelemelerin önemini vurgulamaktadır.

Anahtar Kelimeler: GIS, Mekansal Analiz, Konut Fiyatları, Mekansal Heterojenlik, Türkiye.

**To my beloved family and friends**

## ACKNOWLEDGMENTS

I extend my profound appreciation to Prof. Dr. Zuhal Akyurek, my supervisor, and Dr. Tuncay Kucukpehlivan, my co-supervisor, for their invaluable guidance, unwavering support, insightful advice, constructive criticism, and encouraging perspective during the course of the research.

I also extend my gratitude to Assoc. Prof. Dr. Hediye Tüydeş Yaman for her thoughtful suggestions and comments. Special thanks are due to my classmates, colleagues, and fellow students whose encouragement has been a source of strength during this master's program.

I want to convey my appreciation to Başarsoft Company and its dedicated staff for their cooperation and significant contribution to this research, particularly in providing spatial data.

Heartfelt and sincere gratitude goes to my family my beloved parents and siblings for their unwavering emotional support, encouragement, and understanding during the challenging moments of my academic journey, especially my brothers Yusuf, Mohamed, Abdulrahman and my sister Reema. I also owe a debt of gratitude to my friends who provided support, understanding, and much-needed breaks during the intense periods of this study implementation especially Noor Abu Mousa and Farah El-Silk for their continues encouragement and presence in the toughest moments.

**TABLE OF CONTENTS**

## LIST OF TABLES

## LIST OF FIGURES

## LIST OF ABBREVIATIONS

- UN - United Nations
- GDP - Gross Domestic Product
- OLS - Ordinary Least Squares
- GWR - Geographically Weighted Regression
- GIS - Geographic Information System
- ANN - Artificial Neural Network
- MGWR - Multiscale Geographically Weighted Regression
- ATM - Automated Teller Machine
- OSM – Open Street Map
- SHP– Shape File Spatial Data Type

# CHAPTER 1

# INTRODUCTION

The dynamics of housing prices are essential for understanding urban development and socioeconomic trends. Understanding the spatial distribution of housing prices is important for policymakers, urban planners, and real estate stakeholders to make informed decisions and address disparities in housing affordability and accessibility. According to a report implemented by the United Nations, the world's population is projected to reach 9.7 billion by 2050, with urban areas absorbing much of this growth (URL1). This population surge is expected to exacerbate housing affordability challenges in many urban centers, particularly in developing countries where rapid urbanization is occurring at an unprecedented pace. Research by Saiz (2010) found that population growth in major global cities, such as New York, London, and Tokyo, has led to significant increases in housing prices over time due to increased demand. Moreover, a study by Glaeser et al. (2008) highlights the strong correlation between population growth and housing prices in urban areas, emphasizing the role of supply constraints in driving price escalation. Although the continuation of population growth in the world, the rate of this increase is slowing down, and it is expected that the world will be able to stop growing by the end of the current century and this relationship is illustrated in Figure 1.1.

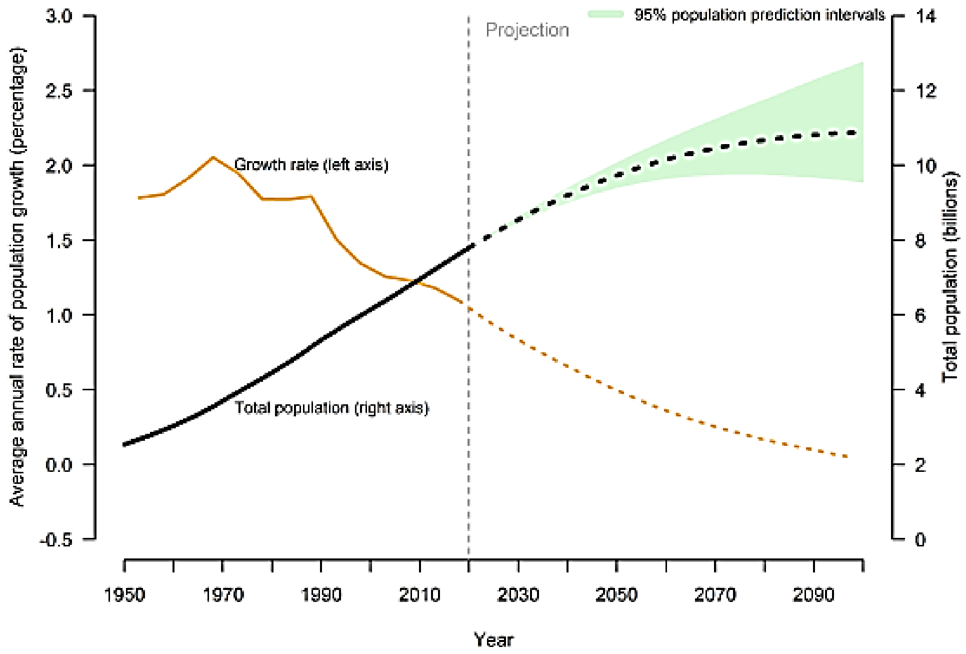



Figure 1.1: Population size along with the annual growth rate for the world: estimates, 1950-2020, and medium-variant projection with 95 per cent prediction intervals, 2020-2100 (URL1).

Turkey, situated at the crossroads of Europe and Asia, has experienced rapid urbanization and population growth in recent decades. The country's population has more than doubled since 1970s, reaching approximately 85 million in 2021 (URL2). This demographic shift has placed immense pressure on housing markets, particularly in major cities like Istanbul, Ankara, and Izmir, where population density is the highest. The research conducted by Habash  and Unanoglu (2020) underscores the significant impact of population growth on housing prices in Turkey (Habash & Unanoglu, 2022). The study found a positive correlation between population increase and housing price appreciation. Figure 1.2 illustarte the population change during the period 2000-2020 in Turkish districts, while Table 1.1 indicates that Ankara is the second highest population city in Turkey with 5.8 million people.

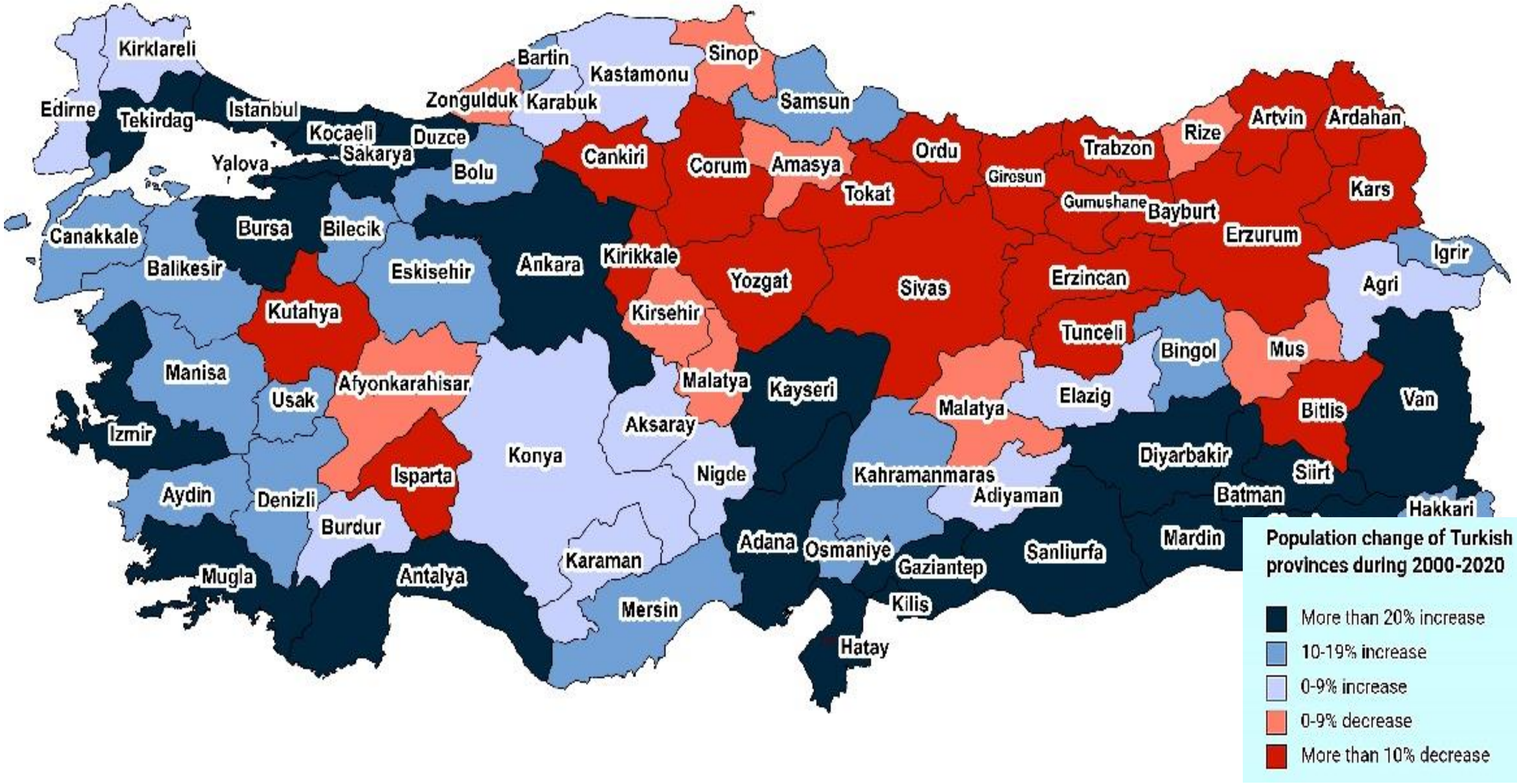



Figure 1.2: Population rate of change in Turkey between the years 2000-2020 (URL3).

Table 1.1: Distribution of the top 5 districts with the highest population, 2023 (URL2).

| City | Population (million) |
|---|---|
| **Istanbul** | 15 655 924 |
| **Ankara** | 5 803 482 |
| **Izmir** | 4 479 525 |
| **Bursa** | 3 214 571 |
| **Antalya** | 2 696 249 |

Residential property values in Ankara have been increasing swiftly in the recent years, making it increasingly difficult for residents to find affordable housing (URL4). The disastrous earthquake happened on February 6, 2023 in Turkey's southeast caused a significant increase in housing prices in Turkey. According to the experts, residents of Istanbul, Izmir, and Ankara are the first to face growing costs of housing for purchase and rental. The residents of the affected areas migrated to large cities. The capital has received many from the earthquake region. Therefore, the rents in Ankara jumped up sharply in the last months. There is also limited research about the spatial distribution of housing transaction in Ankara, which makes it challenging for policymakers and investors to make informed decisions about housing investments and development.

The objective of this study is to examine the elements contributing to the variation in housing prices in Ankara. It is aimed to understand the spatial heterogenity and spatial autocorrelation to identify potential clustering trends in residential property prices within the city. Given the continuous rise in houseing prices, studying these fluctuations becomes necessary due to their significant implications on the city's way of life. Furthermore, it indicates the presence of several influential factors that have contributed to these changes, necessitating a comprehensive examination of their underlying causes. Factors such as the characteristics of the houses like area per square meter, floor number, number of rooms, and proximity to amenities like malls, universities, schools, hospitals, and transportation hubs all play important roles. Larger living spaces, higher floors, and more rooms generally command higher prices due to increased utility and desirability. Additionally, proximity to amenities enhances convenience and quality of life, positively influencing property values. Research has consistently shown the significance of these factors in predicting housing prices, guiding urban planning, investment decisions, and policy formulation (Gyourko et al., 2013; Goodman & Thibodeau, 1998). By considering these dynamics, a more comprehensive understanding of housing market in Ankara could be achieved, leading to informed decision-making and policy formulation. For this purpose, a GIS application was developed, incorporating various analysis tools to gain a deeper understanding of housing transaction patterns within the Ankara metropolitan area and to assess the impact of several amenity factors on these prices.

The structure of the study is organized as follows: Chapter 2 reviews recent and relevant literature. Chapter 3 introduces the study area, the data utilized, and the methodology applied. Chapter 4 explaines the developed GIS tool. Chapter 5 presents the results of the analysis. Chapter 6 presents the discussions and Chapter 7 offers the conclusions.

# CHAPTER 2

# LITERATURE REVIEW

Spatial Analysis have been implemented on housing prices for several decades, since the spatial location of a house plays a pivotal role in determining its accessibility to many factors. Therefore, many academic researches pointed to the importance of such study (Can, 1998).

Many researches implemented spatial analysis on the subject of residental and urban properties prices. Leung et al. (2006) investigated empirically the degree of price dispersion within Hong Kong real estate market during the 1990s as it is considered a “thick” market (a market that exhibits a substantial level of trading activity). While Raslanas et al. (2006) implemented spatial analysis to apply a comparison on differences in housing prices for both areas of two-bedroom apartments in South-eastern London and three-room flats of comparable size in Vilnius by applying statistics methods for the analysis process. Additionlly, Baker et al., (2016) employed well-established methods from the field of population geography to evaluate the influence of Australia's growing housing affordability crisis on the spatial distribution of the country's population across regions with varying levels of advantage. Bagheri and Shaykh-Baygloo, (2021) investigated the existence of the relationship between the urban growth and the housing prices of Isfahan metropolitan region using spatial analysis methodologies.

Besides spatial factors, some researches included the temporal factor in their analysis as an importatant element in understanding the change in housing prices. For example, spatial temporal iteration of the regression models used by Crespo et al. (2007) inorder to facilitate the prediction and eventual interpolation of local parameters over time.

Altuzarra and Esteban (2011) urged that the recent surge in home prices globally has sparked a fresh interest in studying both the long-term and short-term factors influencing housing prices. Therefore, Altuzarra and Esteban (2011) investigated the Granger causation relationship between the land prices and the prices of residential properties in Spain. They used a statistical concept to determine whether one time series variable can predict another time series variable. The analysis focused on quarterly observations for both housing and land price per square meter from provinces in Spain throughout the period of 2005-Q1 to 2010-Q2.

Barros et al. (2013) conducted the spatial analysis methodologies to investigate the temporal pattern of residential property prices across 69 Chinese cities between 2005-2010. A study in China used big data amount from online real estate platforms to analyze housing price disparities spatially and temporally, emphasizing the significance of big data analytics (Li et al., 2017). Another study in Malaysia applied the temporal analysis to explore the impact of macroeconomic factors on housing transactions and demand from 2002 to 2015 (Kok et al., 2018). Same approch was also implemented on Turkey's housing properties, when Akkay (2021) examined the correlation between housing prices and macroeconomic determinants within the Turkish economy from 2010 to 2020, and based on the outcomes, it was deduced that the capacity of macroeconomic determinants to elucidate the fluctuations in housing prices is influenced by countryspecific factors.

De Bruyne and Van Hove (2013) emphasized the importance of considering geographical location, alongside socioeconomic variables and real estate characteristics, in understanding housing price variations in Belgium. The results indicated that distance and travel time to economic centers significantly impact housing prices.

Incorporating geographical barriers into the model provides a unique perspective on housing price dynamics. However, the literature review shows that, in addition to studies exploring determinants of housing prices according to geographical location, there are others that have examined more factors and their contribution to the price distribution. From these studies, Kestens et al. (2006) integrated household-level data such as category, age, level of education and earnings, into the Geographically Weighted Regression (GWR) model to understand urban dynamics and socioeconomic conditions influencing housing prices in Quebec City, Canada, from 1993 to 2001. While, Hu et al. (2016) used an index system to analyze the spatial correlation between land price and relevant factors in Wuhan, China. Categorized twelve factors into land characteristics, locational elements, and neighborhood attributes, aiming to identify spatial patterns and associations within the city.

Additionally, applying both the temporal factor along with macroeconomic factors influencing housing prices; Ghaderi and Izady (2016) used regression models to study these factor’s influence on housing prices in Iran from 1972 to 2013. They indicated that urbanization pace, personal earnings, rental costs, and Gross Domestic Product (GDP) which is a measure of the total value of all goods and services produced within a country's borders over a specific period of time, typically annually or quarterly; to be primary influencers, with construction costs playing a secondary role. Similarly, Wang et al. (2019) investigated housing price influences in Taitung city from 2013 to 2017, in which regression models and spatial aggregation were used.

Moran’s I was used in examining spatial effects on the housing market and their relationship to the crime rates in a study in the Belfast Metropolitan Area (UK). Using Moran's I model, comprehensive dataset including property characteristics, neighborhood factors, socioeconomic indicators, and crime rates were analyzed. The study showed no significant spatial correlation, indicated by low and negative Moran's I values, suggesting a random arrangement of variables (McIlhatton et al., 2016).

In a study performed on residential land prices in Kuwait, Moran's I analysis identified spatial clustering in land prices across Kuwait in which conventional and Bayesian spatial regression methods were used to analyze factors affecting prices. (Mostafa, 2018). Souza et al. (2021) investigated spatial autocorrelation in mean housing prices in Salvador, Brazil. The Global Moran's I Index was calculated, revealing significant positive spatial autocorrelation among prices across the city's 163 districts. This suggests a varied distribution pattern of average housing prices throughout Salvador's urban landscape.

The Ordinary Least Squares (OLS) model and Geographically Weighted Regression (GWR) method are commonly used to analyze the relationship between housing prices and various factors. OLS is a widely recognized regression technique, providing a global model to understand and predict variables. It captures overall dynamics by formulating a single regression equation (URL5)  On the other hand, GWR extends OLS by considering spatial variations in relationships, addressing spatial autocorrelation. GWR builds separate equations for each location, incorporating nearby data to uncover spatial dynamics (URL6).

Subsequent researches explore both regression methodologies. For example,  Basu and Thibodeau (1998) investigated spatial autocorrelation in single-family property transaction prices in Dallas, Texas. They used a semilogarithmic housing price equation and a spherical autocorrelation function with data from over 5000 transactions. Comparing OLS and krigged Estimated Generalized Least Squares (EGLS) methods in predicting property values revealed spatial autocorrelation within submarkets. Krigged EGLS outperformed OLS predictions in six out of eight submarkets, except where residuals showed no spatial correlation and the estimated semivariogram had a large variance.

Sun et al. (2005) analyzed 54,282 condominium transactions in Singapore (1990-1999) using a second-order spatio-temporal autoregressive model. This model addressed spatio-temporal autocorrelations and heteroscedasticity, distinguishing spatial effects into building and neighborhood components. In this study, Bayesian estimation proved more robust than conventional OLS, revealing distinct behaviors in building-specific price indices compared to aggregate market indices.

In a study where Geographically Weighted Regression (GWR) was used to segment the housing market, utilizing data from the Italian Ministry of Economy and Finance;

GIS's ability was demonstrated to identify price clusters, favoring GWR over Ordinary Least Squares (OLS) based on residual and Akaike Information Criterion (AIC) analysis. GWR revealed high housing prices not only in urban but also agricultural areas, showcasing the complex dynamics influencing housing prices beyond urban regions (Manganelli et al., 2014).

In order to enhance the geographically weighted regression (GWR) model, Huang et al. (2010) added temporal effects, creating a model called regression with geographical and temporal weighting (GTWR). This improved model analyzes real estate market data by addressing both spatial and temporal variability simultaneously.

Moreover, Cao et al. (2019) examined spatial price fluctuations in Singapore's public housing resale market, where ordinary least squares (OLS), geographically weighted regression (GWR) using Euclidean distance, and GWR based on travel time methods were compared. The results showed both GWR models outperform OLS, with the travel time-based GWR demonstrating superior fit compared to the Euclidean distance-based GWR.

Other researches are comparing OLS with Artificial Neural Network (ANN), Mankad (2022) found that using Artificial Neural Network (ANN) improves housing price estimations compared to Ordinary Least Squares (OLS) in Vadodara city, India. Integrating structural and spatial attributes, ANN overcomes limitations like nonlinearity and multicollinearity seen in the traditional hedonic model based on OLS.

Apaydin and Güneş (2022) used Geographically Weighted Regression (GWR) to explore spatial variation in Ankara's metropolitan housing market. GWR, employing a Gaussian kernel weighting function with adaptive bandwidth determined through cross-validation, outperformed conventional ordinary least squares regression, capturing spatial effects better. The analysis revealed spatial variation in the housing market, indicating non-stationarity with varying explanatory power and parameter estimations across the area. The study sets the stage for examining changes in housing prices and their impact on regression modeling and analysis in Ankara's metropolitan area.

Moreover on comparing different regression models, Tomal and Helbich (2023) compared the performance of geographically weighted quantile regression with spatial autoregressive errors (GWQR-SAR) and other regression models using rental price data from Warsaw, Poland, and Amsterdam, The Netherlands. They found that GWQR-SAR outperformed alternative models, providing better accuracy, especially at the extremes of the dependent variable distribution.

Liu and Strobl (2023) examined resale prices of houses and neighborhood characteristics in Zhuhai, China by using OLS, GWR, and Multiscale Geographically Weighted Regression MGWR. They found that GWR and MGWR effectively captured local spatial variations, outperforming OLS in fit quality and parameter estimates.

# CHAPTER 3

# STUDY AREA AND METHODOLOGY

## 3.1 Study Area And Data Used In This Study

Housing prices are influenced by various factors, resulting in fluctuations in different regions. Among these factors, geographical location holds significant importance in determining housing prices. Academically, statistical tests are employed to assess the impact of other factors on housing prices. However, it is important to emphasize that the factors affecting housing prices may vary across different study areas. Therefore, it is not possible to attribute a single factor as the sole determinant of house transactions universally. Analyzing the impact of spatial factors on housing prices is of utmost importance when it comes to comprehending the housing sector within a specific area. By doing so, one can evaluate how it impacts the local economy, housing density, population distribution, and urban development. This analysis also sheds light on the availability of important amenities like education, healthcare, shopping facilities, and transportation infrastructure. Consequently, such studies offer invaluable insights into the housing market, facilitating informed decision-making processes. Moreover, utilizing the findings of spatial analysis within applications that incorporate interactive maps presents an opportunity to understand the spatial distribution of various housing prices and their correlation with other factors in a practical and interactive manner. This approach enables a deeper understanding of the clustering patterns that emerge from different inputs, ultimately aiding in the prediction of future forms of spatial distribution for housing prices in the region.

Ankara Metropolitan Area, chosen as the study area for its dynamic urban landscape and significant population growth, serves as a compelling backdrop for investigating the spatial dynamics of housing prices, contributing valuable insights to urban development and policy planning. Figure 3.1 describes how the populate rate in Ankara metropolitan area changed between the years 2019-2023. Figure 3.2 depicts the study area addressed in this study.

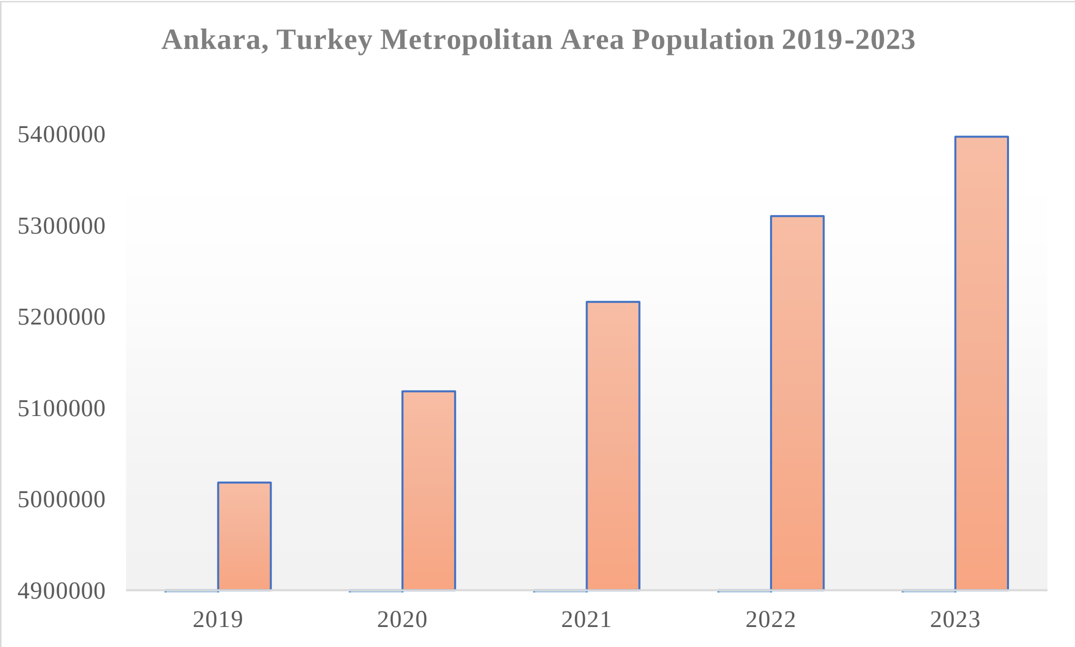



Figure 3.1: Ankara, Turkey Metropolitan Area Population 1950-2023 (URL7).

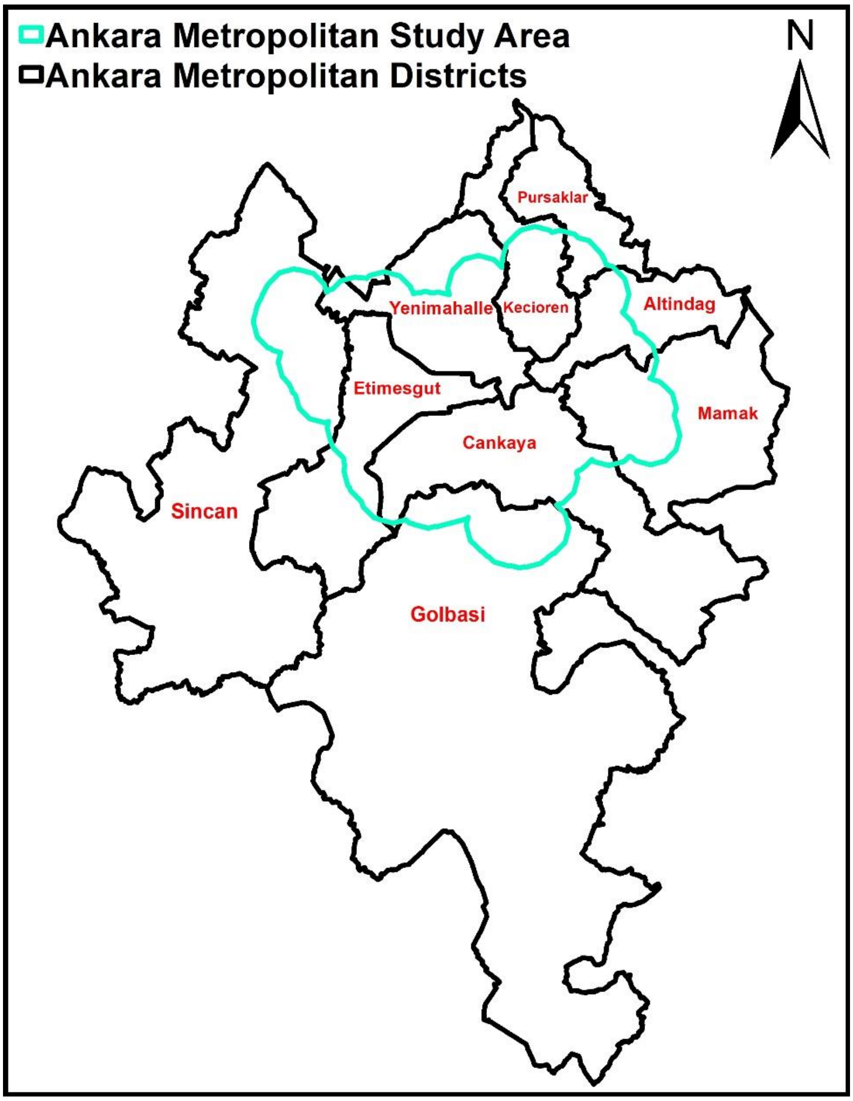



Figure 3.2: Ankara Metroplitan Study area.

Over 45,000 data records from 24 districts in Ankara were collected. The data obtained from real estate portals like Sahibinden, focused on one neighborhood per district and was gathered in the last quarter of 2022 (URL8). The housing transactions was originally in Turkish Lira (TL) and was converted to United States Dollars (USD) using the exchange rate in the last quarter of 2022, where ($1= 18.6123 TL) (URL14). The data extraction process was time- and effort-intensive, as the information was initially obtained in JSON format, requiring conversion to

Excel tables before being further processed into shapefiles for GIS analysis. Figure 3.3 shows a template from the website main interface and how the data of housing transaction are distributed within Ahlatlibel neighbor in Çanakaya district.

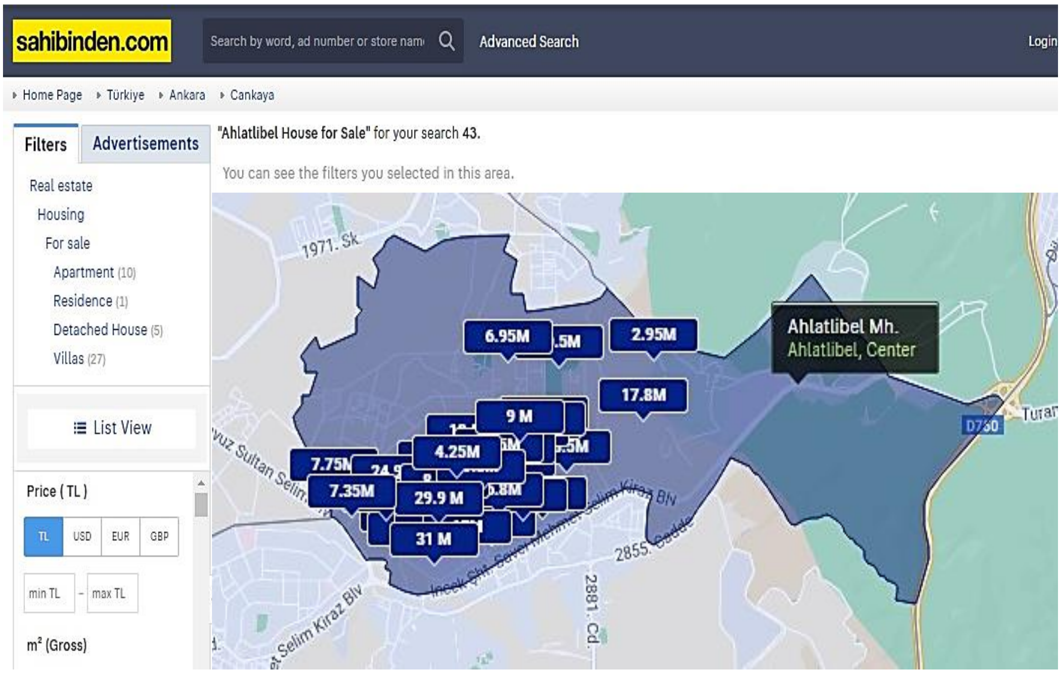



Figure 3.3: Distribution of housing units within the Ahlatlibel neighborhood located in Çankaya district as shown in Sahibinden real estate website in August, 2024 (URL8).

Data cleaning involved the elimination of records with missing or incomplete information, ensuring that only reliable and usable data were retained for analysis. After processing, the data was organized in a table format with necessary attributes using Excel before being uploaded into GIS applications for further analysis. These transactions are concentratedly distributed at the center of the Ankara metropolitan area (Figure 3.4 (a)). In Figure 3.4 (b) the amount of these transaction in USD ranging from lowest prices represented in blue color to highest prices represented in red color. Most of the high-priced transactions, represented by the red dots, are distributed in the center-west of Çankaya district, which is considered one of the most important districts in Ankara due to the presence of many economic, educational and health facilities there.

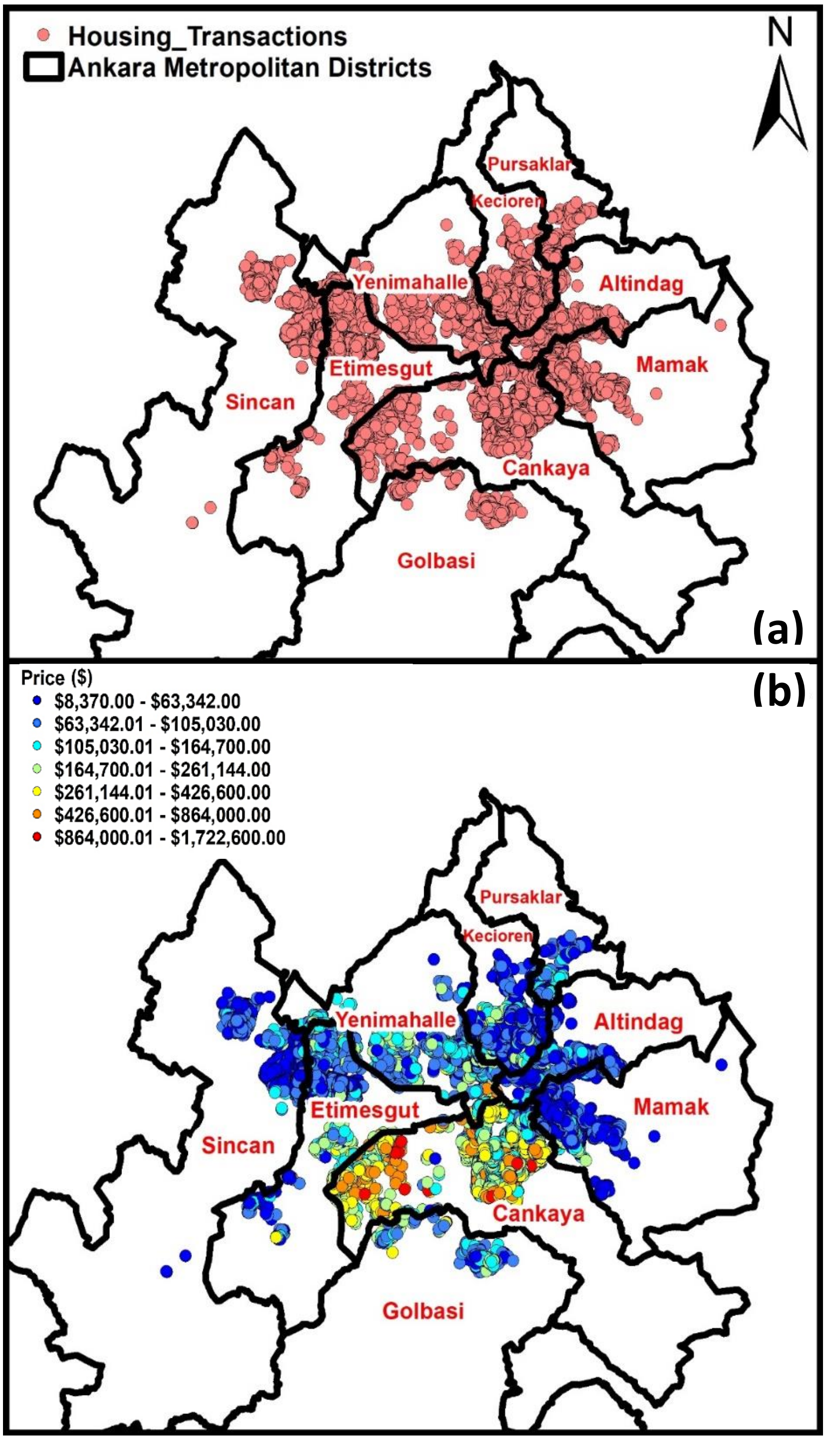



Figure 3.4: (a) Distribution of housing units within the Ankara metropolitan area in the last quarter of 2022 (b) Point density map of prices in USD obtained from (URL8) real estate website.

Sample from the main dataset used in the spatial analysis in this study for randomly selected locations are mentioned in Table 3.1.

Table 3.1: Sample from housing transaction data.

| **Attributes** | **Location1** | **Location2** | **Location 3** | **Location 4** |
|---|---|---|---|---|
| **Price in (USD)** | **108,000** | **79,920** | **81,000** | **175,500** |
| **Neighbor** | Atifbey | Atifbey | Aydinlikevler | Aydinlikevler |
| **District** | Altindag | Altindag | Altindag | Altindag |
| **Area / (m2)** | 135 | 135 | 130 | 195 |
| **Room Number** | 4 | 4 | 4 | 5 |
| **Floor Number** | 7 | 3 | 3 | 3 |
| **Dis. to University (m)** | 1403.4 | 1387.8 | 2908.2 | 2462.4 |
| **Dis. to Hospital (m)** | 375.8 | 432 | 110.7 | 483.7 |
| **Dis. to Mall (m)** | 2673.6 | 2661.8 | 1616.7 | 1570.4 |
| **Dis. to Transport (m)** | 145 | 124.5 | 85 | 95.1 |
| **Dis. to School (m)** | 119.02 | 178.24 | 192.24 | 319.4 |
| **Dis. to ATM (m)** | 48.4 | 72.8 | 121.5 | 29 |
| **Dis. to Bank (m)** | 48.4 | 73.1 | 126.4 | 36.2 |
| **Dis. to BIM &A101 (m)** | 90.6 | 79.3 | 43 | 85.1 |
| **Dis. to Migros (m)** | 1329.2 | 1380.6 | 335.1 | 114.7 |

Initially, more than 45 000 values of house transfers were collected, each with a set of information necessary to do spatial analysis, and after studying this data, it was found that there was a loss in one or more of these data during some conversions, which were excluded from the study, so that the number of data became 42 000. In a recent study, it is found that in Ankara metropolitan area, property sales transactions accounted for over 80 percent of the total comprising eight districts in 2021 (Apaydin & Güneş, 2022). As a preprocessing this study found that 96.51% of the housing prices are located in the Ankara metropolitan area consisting of nine districts, same as the findings of Apaydin and Güneş (2022), only additionally, Gölbaşı district was found to be significant to consider since that the percent of transaction for that district is more than 3% as shown in Table 3.2. Therefore, it is more convenient to work on this particular area in Ankara in terms of conducting regression modeling and autocorrelation analysis than considering the whole city.

Table 3.2: a) Distribution percent of real estate transaction in Ankara, 2022 b) Distribution percent of real estate transaction in Ankara,2021 used in (Apaydin & Güneş, 2022).

a)

| **District** | **N** | **%** |
|---|---|---|
| Kecioren | 7777 | 18.34 |
| Çankaya | 7127 | 16.81 |
| Etimesgut | 5360 | 12.64 |
| Sincan | 5355 | 12.63 |
| Mamak | 4679 | 11.04 |
| Yenimahalle | 4000 | 9.44 |
| Altindag | 2744 | 6.47 |
| Pursaklar | 2399 | 5.66 |
| Golbasi | 1476 | 3.48 |
| **Total** | **40917** | **96.51** |

b)

| **District** | **%** |
|---|---|
| Kecioren | 12.97 |
| Çankaya | 15.57 |
| Etimesgut | 10.97 |
| Sincan | 10.06 |
| Mamak | 11.62 |
| Yenimahalle | 10.33 |
| Altindag | 7.54 |
| Pursaklar | 3.38 |
| **Total** | **82.43** |

Since 96.51% of household transactions are located in the metroplolition area of Ankara, the districts located within this region were adopted in the analysis presented by this study, as the number of these transactions reached 40,910, and it is the final data on which all subsequent analysis in this study were conducted. This final dataset includes the following attributes: Housing prices in Turkish Lira, Area/m2, Room Number. and floor Number. which were obtained from real estate website. In addition to amenities spatial data for transportation stations, universities, schools, shopping malls and hospitals, were obtained from Başarsoft company (URL9).

In Apaydin & Güneş (2022) factors such as the age of the building, and distance to police stations were also used in addition to the factors considered in this study. The selection of the eight key explanatory attributes in the housing price spatial analysis is grounded in their critical influence on property values and urban dynamics. Area per m2, floor number, and number of rooms capture fundamental aspects of housing characteristics (Nistor & Reianu, 2018). The inclusion of distance to transportation, malls, hospitals, schools and universities reflects their recognized impact on property prices and desirability (Hewitt & Hewitt, 2012). These factors collectively provide a comprehensive understanding of the multifaceted determinants shaping housing prices within the Ankara Metropolitan Area, therefore this study sticked with them and did not use he other factors used by Apaydin & Güneş (2022). Amenities factors are spatially distributed within the Ankara metropolitan area as presented in Figures 3.5 - 3.16 . The figures show in order how bus stations, metro stations, railway stations, universities, hospitals, malls, Automated Teller Machine (ATM), banks, budget supermarkets(BIM&A101), supermarkets (Migros) and schools are distributed within the Ankara metropolitan area.

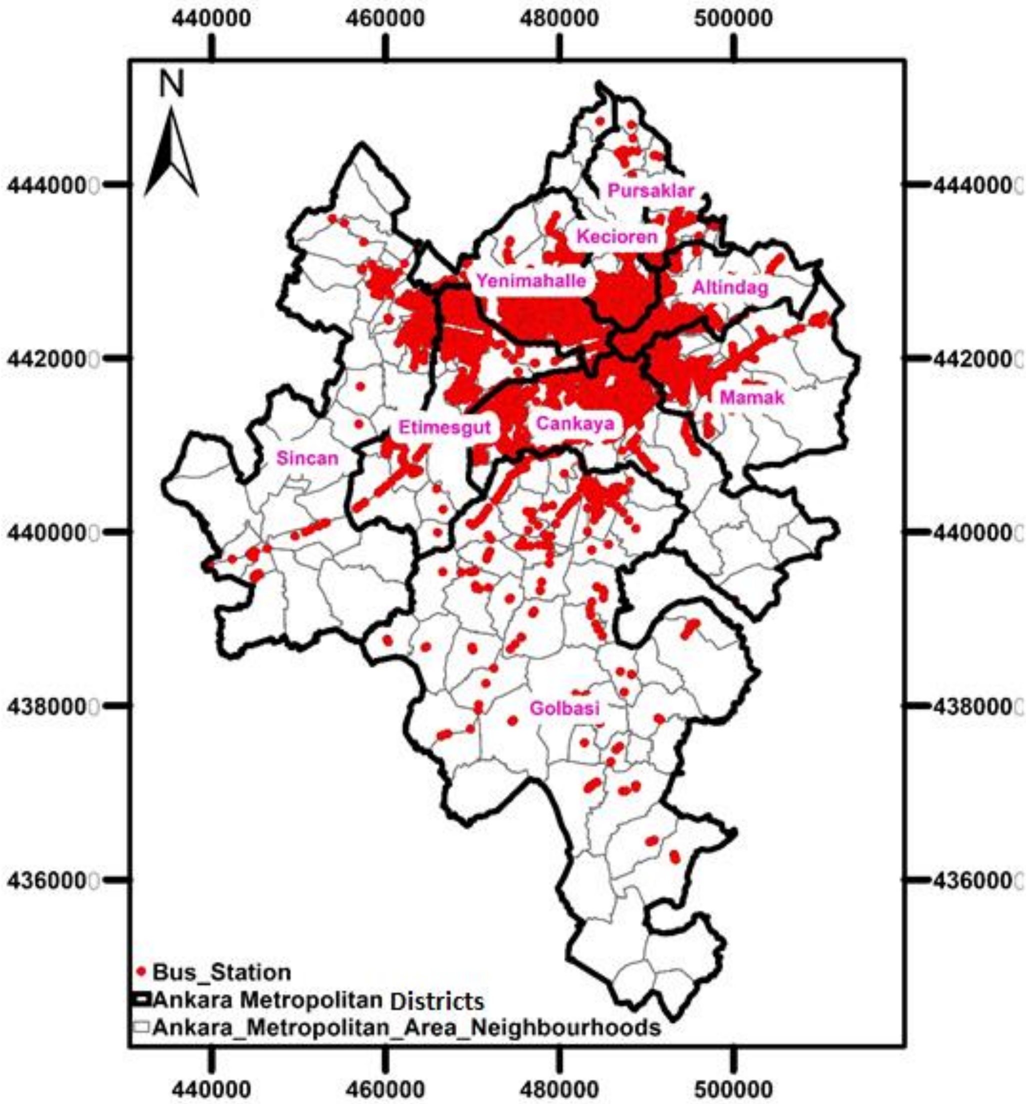



Figure 3.5: Spatial distribution of bus stations within the Ankara metropolitan area.

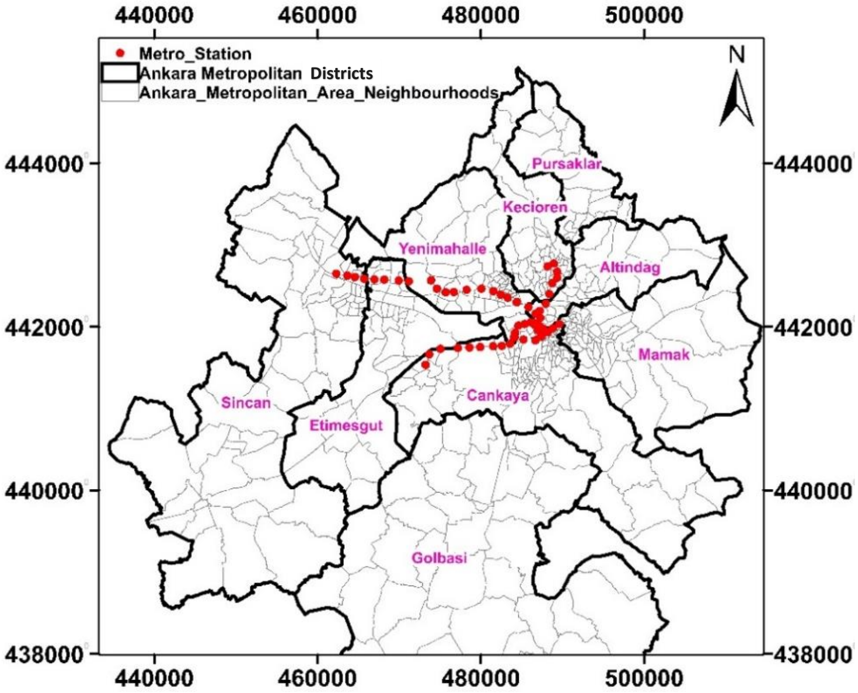



Figure 3.6: Spatial distribution of metro stations within the Ankara metropolitan area.

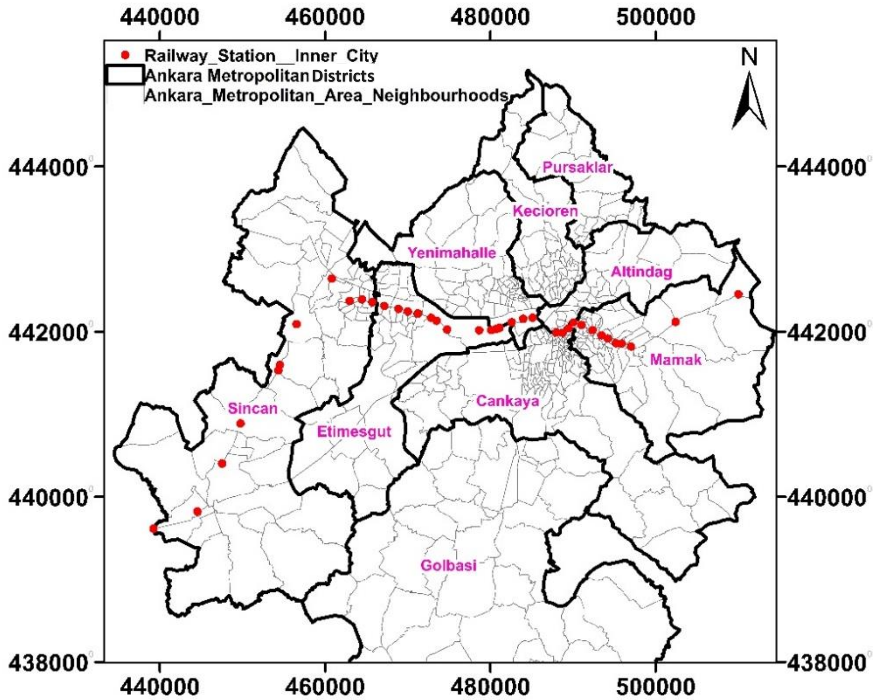



Figure 3.7: Spatial distribution of railway stations within the Ankara metropolitan area.

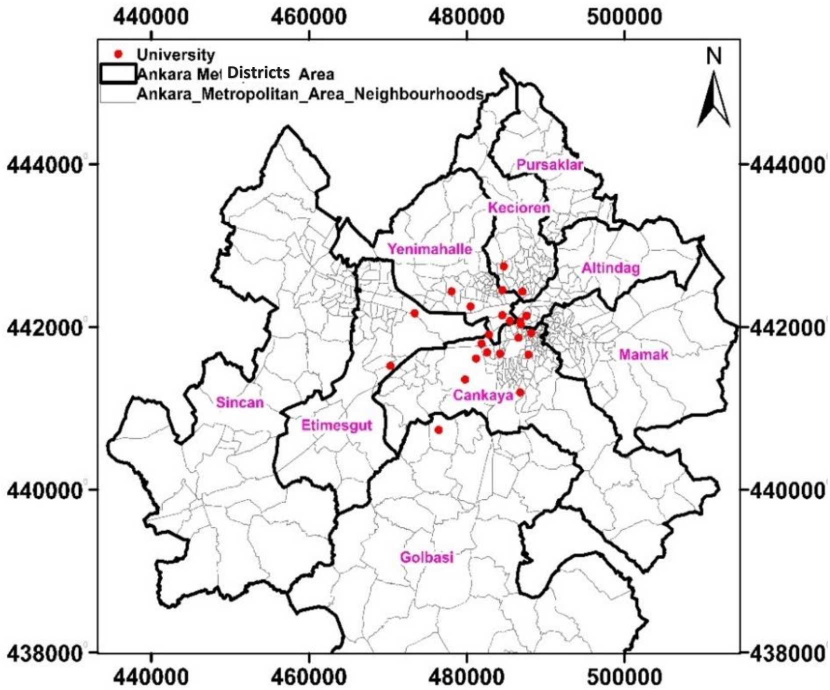



Figure 3.8: Spatial distribution of universities within the Ankara metropolitan area.

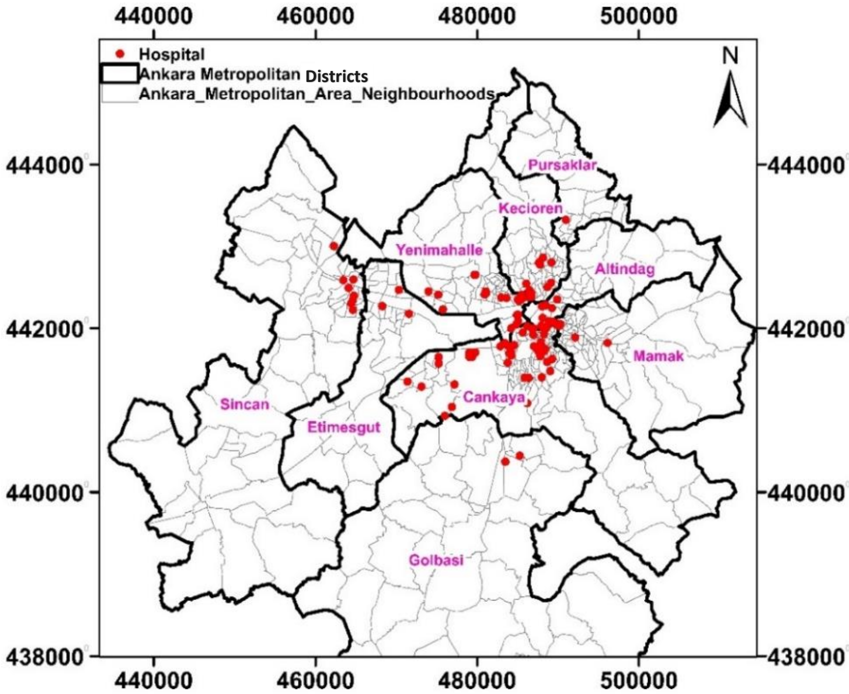



Figure 3.9: Spatial distribution of hospitals within the Ankara metropolitan area.

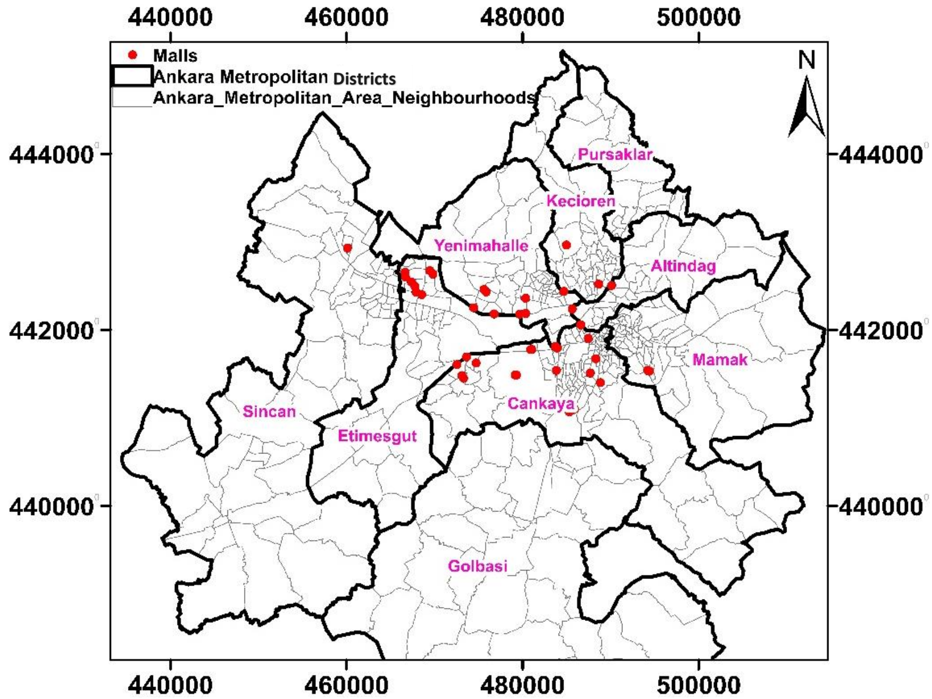



Figure 3.10: Spatial distribution of malls within the Ankara metropolitan area.

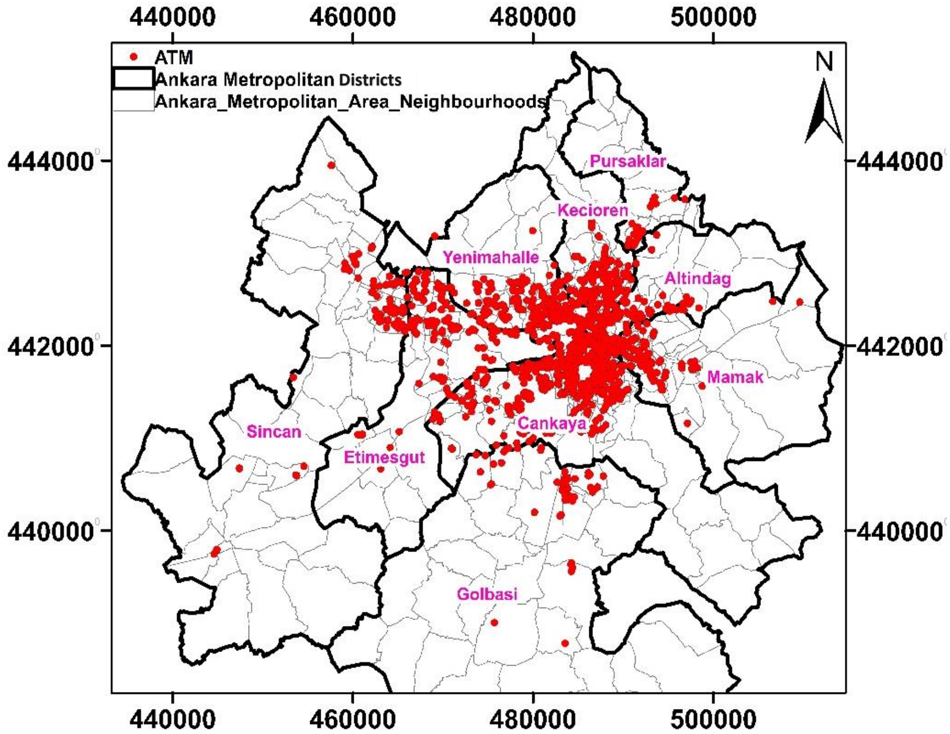



Figure 3.11: Spatial distribution of  ATM’s within the Ankara metropolitan area.

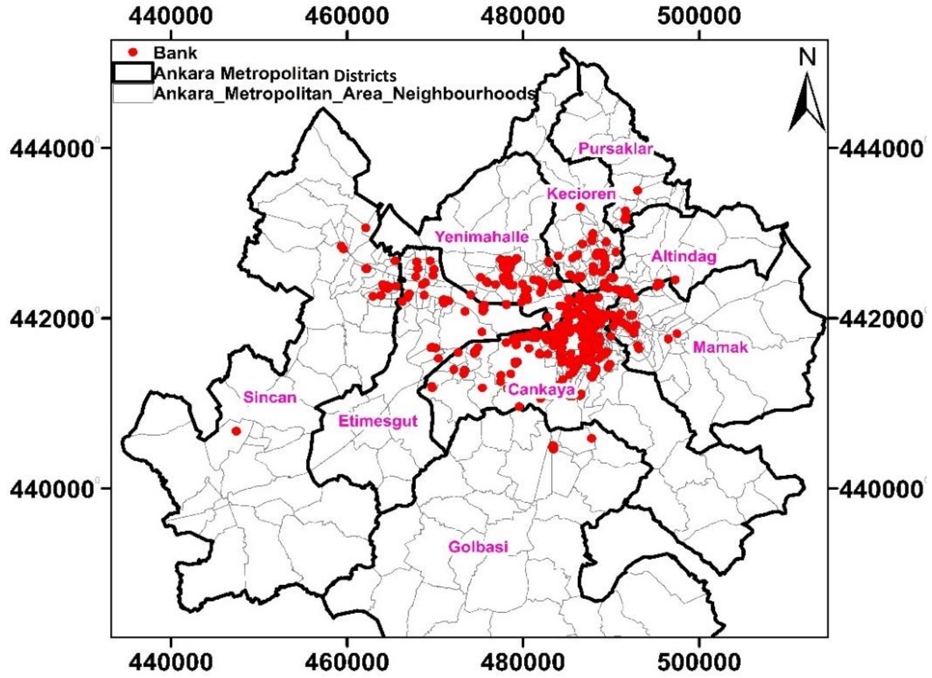



Figure 3.12: Spatial distribution of banks within the Ankara metropolitan area.

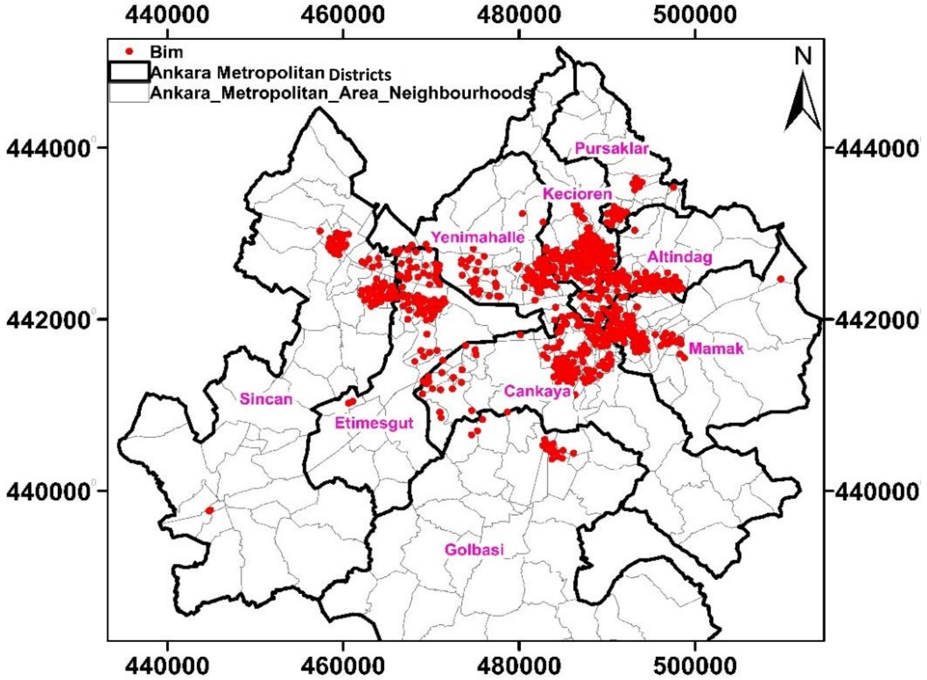



Figure 3.13: Spatial distribution of BIM Budget markets within the Ankara metropolitan area.

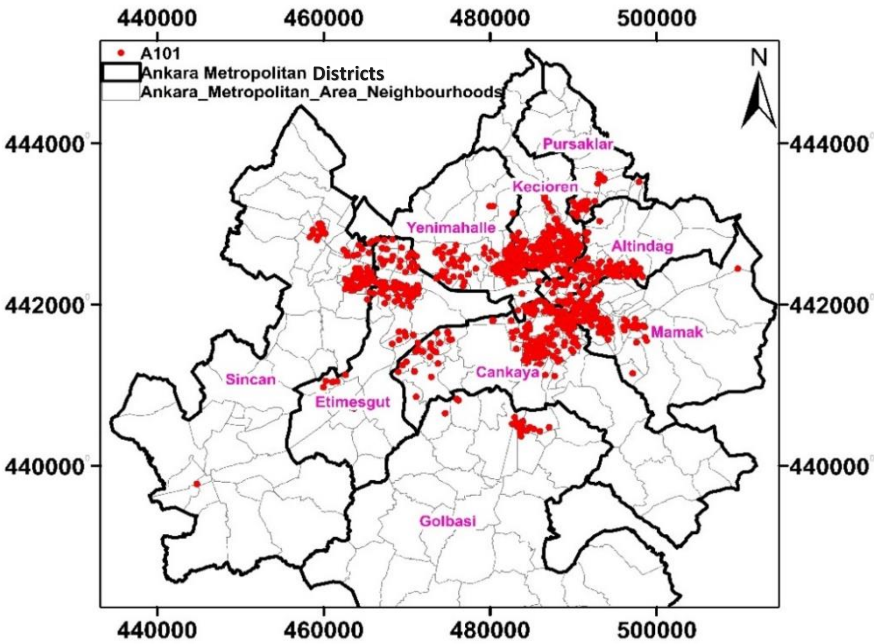



Figure 3.14: Spatial distribution of A101 Budget markets within the Ankara metropolitan area.

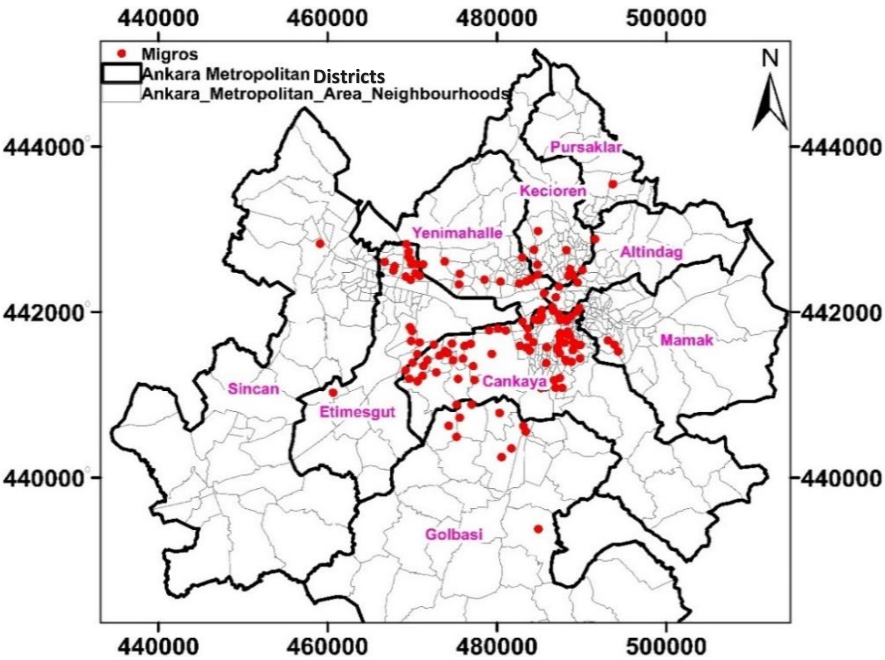



Figure 3.15: Spatial distribution of Migros Super market within the Ankara metropolitan area.

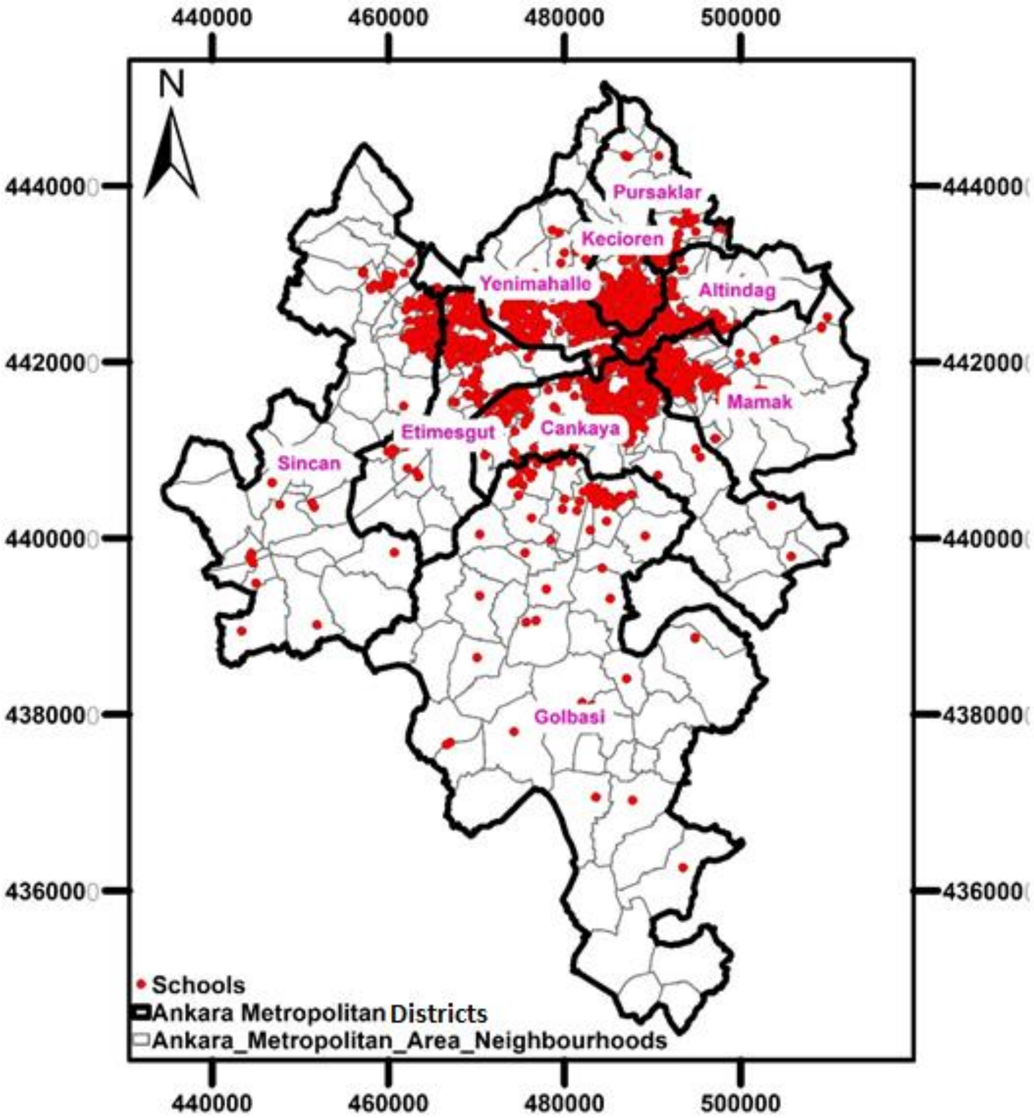



Figure 3.16: Spatial distribution of schools within the Ankara metropolitan area.

## 3.2 Data Preprocessing

The ameneties factors obtained from Başarsoft company were later preprocessed to compute the shortest distance between each feature of these amenities to the closest house using the Euclidean distance method which is a metric that quantify the closeness between two points in the feature space (Singh, 2024). Figure 3.17 visually demonstrates the calculation of Euclidean distance. In a two dimensional space, the Euclidean distance between any random two points (a1,b1) and (a2,b2) would be calculated according to the following equation, where d is the Euclidean distance between two points, a1 and b1 are the x- and y-coordinates of the first point, a2 and b2 are the x- and y-coordinates of the second point.

$$d = \sqrt{((a2 - a1)2 + (b2 - b1)2)} \quad (3.1)$$

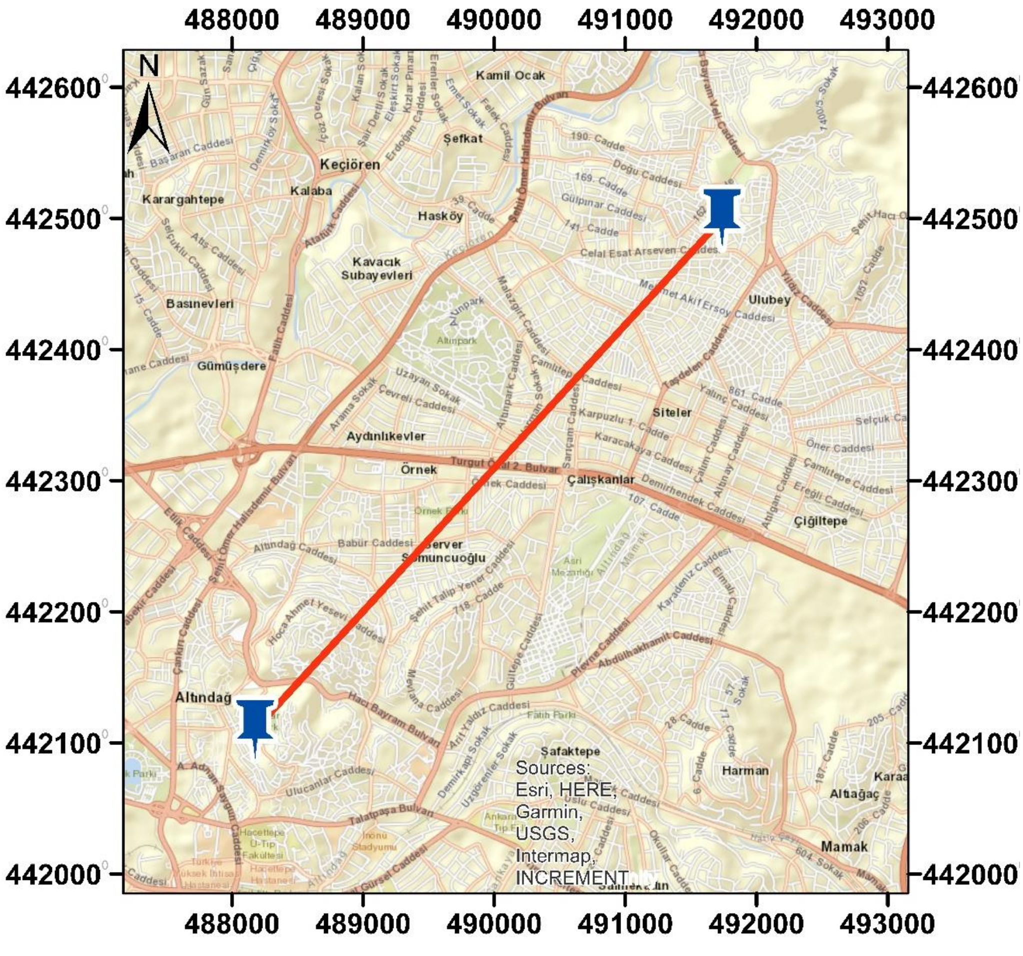



Figure 3.17: Eucliden distance.

This method can be applied under the proximity toolset in GIS that comprises tools utilized for assessing the closeness of features within a single feature class or between two feature classes.

Besides the Euclidean distance calculation, the proximity of the housing transaction to the amenities is calculated. In Figure 3.18 the question mark label represents one of the housing transaction features, yellow circles and purple triangles represent two of the four classes of the ameneties factors, for example if yellow circles were the universities features, proximity tool computes the distance between the targeted housing transaction question mark points and all the universities located features to detarmain at the end the shortest distance between this housing transaction and all features of the universities class, the same methodology is applied to all other housing transactions and the rest of the ameneties factors.

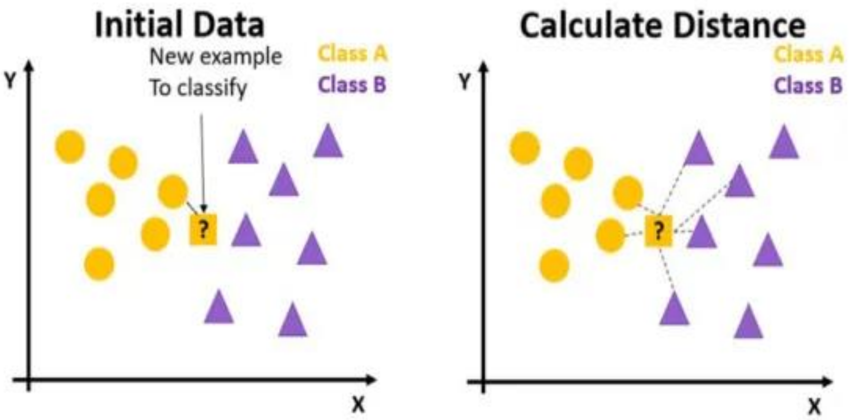



Figure 3.18: Proximity methodology used to measure distance between two or more different features (Singh, 2024).

To provide a comprehensive understanding of the variables employed in this analysis, Table 3.3 presents the statistical summaries of each variable, where their characteristics are given in Appendix A. This table includes key descriptive statistics, such as the first quartile (Q1), the third quartile (Q3), the mean, standard deviation, minimum, and maximum values, which offer insights into the central tendencies and dispersion of the data. The first and third quartiles are statistical measures used to describe the distribution of the data set. The first quartile (Q1) is the value below which 25% of the data falls, representing the lower quarter of the dataset. The third quartile (Q3) is the value below which 75% of the data falls, marking the boundary for the upper quarter.

These summaries are necessary for interpreting the subsequent analytical results and understanding the underlying characteristics of the dataset.

Table 3.3: Statistical Summaries (N= 40910).

| Variable | Mean | Std. Dev. | Min | Q1 | Median | Q3 | Max |
|---|---|---|---|---|---|---|---|
| Price ($) | 98565.81 | 81032.74 | 8370 | 51300 | 75546 | 118530 | 1722600 |
| Area/m2 | 148.61 | 96.62 | 20 | 120 | 135 | 165 | 10000 |
| Room Number | 4.09 | 0.94 | 1 | 4 | 4 | 4 | 12 |
| Floor Number | 3.41 | 4.46 | -4 | 1 | 2 | 4 | 30 |
| dist. to Transport(m) | 177.65 | 223.50 | 8.16 | 83.0 | 134.32 | 204.93 | 13272.69 |
| dist. to University(m) | 5952.71 | 7303.55 | 13.75 | 2019.9 | 4164.73 | 7450.99 | 119756.51 |
| dist.to Mall(m) | 3765.77 | 6029.84 | 7.37 | 1533.8 | 2387.11 | 3753.20 | 119048.73 |
| dist.to Hospital(m) | 1967.42 | 1865.47 | 4.49 | 873.2 | 1495.07 | 2551.76 | 22335.18 |
| dist. to School(m) | 279.01 | 206.34 | 2.68 | 146.6 | 234.71 | 348.65 | 2297.53 |
| dist.to ATM(m) | 374.50 | 322.97 | 0.49 | 167.5 | 305.29 | 489.62 | 3770.22 |
| dist.to Bank(m) | 944.13 | 1018.77 | 1.93 | 403.9 | 724.39 | 1150.86 | 12031.68 |
| dist.to Supermarket(m) | 1759.59 | 1554.38 | 1.15 | 611.1 | 1285.01 | 2480.40 | 17369.85 |
| dist.to Budget Supermarket(m) | 293.46 | 519.97 | 0.35 | 112.97 | 187.63 | 309.61 | 8625.80 |

Histograms were generated in order to identify potential outliers, assess data quality, and check for adherence to statistical assumptions critical for spatial analyses. The histogram are implemented using the Natural Break method with seven classes for the frequency distribution of each variable. Histograms are shown in Appendix B subsection.

## 3.3 Proposed Methodology

### 3.3.1 Methodology workflow

The methods used in this study include data preparation, empirical analysis, spatial analysis, spatial modeling, visualization and application implementation. Figure 3.19 illustrates a flow chart depicting the interrelation of the methods.

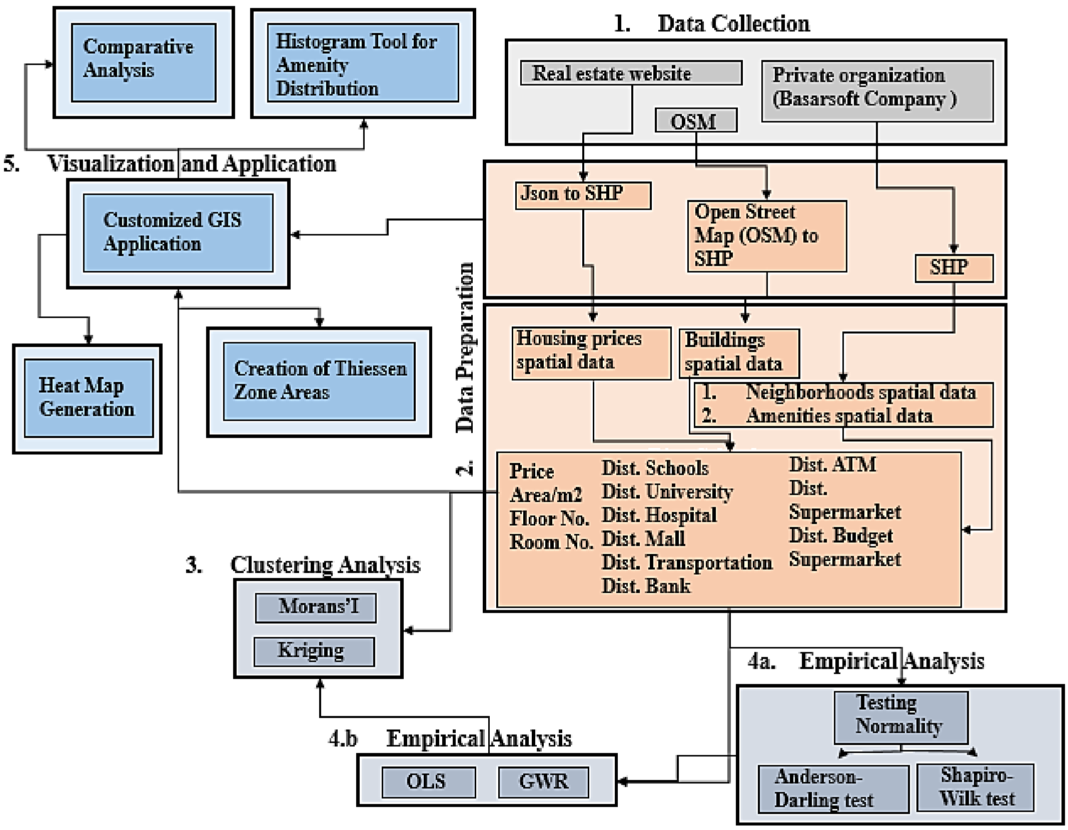



Figure 3.19: The flow chart of the methodology.

### 3.3.2 Clustering analysis

In order to understand the heterogeneity of housing prices a clustering analysis is important to be applied. Clustering is basiclly a grouping mechanism in whish each set of data records that are similar in some sense are gathered in one group together.

This study focuses on determining the correlation within the dataset based on property prices. The clustering algorithm relies on the spatial distribution of price information to segregate the dataset. Morans' I autocorrelation is employed as the clustering analysis tool in this study.

#### i Morans'I

Moran's I measures the correlation between neighboring observations in a pattern, ranging from -1 to 1. Positive values indicate spatial clustering, while negative values suggest dispersion. A result near zero shows no spatial autocorrelation (URL10). The tool calculates a z-score and p-value to assess the statistical significance, with a small p-value indicating that the observed pattern is unlikely to be random, leading to the rejection of the null hypothesis of random distribution (CSR - Complete Spatial Randomness).

A high positive z-score signifies clustering of similar values, while a negative z-score indicates dispersion. Moran's I was applied to the housing price dataset using ArcGIS, with inverse distance used as the spatial relationship model, which accounts for distance decay. The Euclidean distance method was chosen as the distance band for continuous data, like housing price changes. To define the threshold distance, a Kriging analysis helps determine at what distance housing prices cease to be correlated.

#### ii Kriging

Kriging is a geostatistical technique that estimates a surface from dispersed points with associated z-values, generating a Semivariogram (URL11). This model identifies spatial correlations, assuming that the distance or direction between points reveals surface variation. The Kriging tool fits a mathematical function to nearby points, reaching the range when the model flattens, beyond which points no longer exhibit spatial autocorrelation. The semivariogram reaches a value at the range known as the sill, and the partial sill is calculated by subtracting the nugget from the sill. Lags represent the distances between pairs where the variogram is calculated, and the nugget effect occurs at very small distances, indicating variation when the semivariogram value is greater than 0 (URL11). For instance, if the semivariogram intersects the y-axis at 2, the nugget is 2. Ordinary Kriging was implemented to create a Prediction Surface, which predicts housing prices based on spatially continuous data (URL12). Figure 3.20 summarizes the semivariogram components.

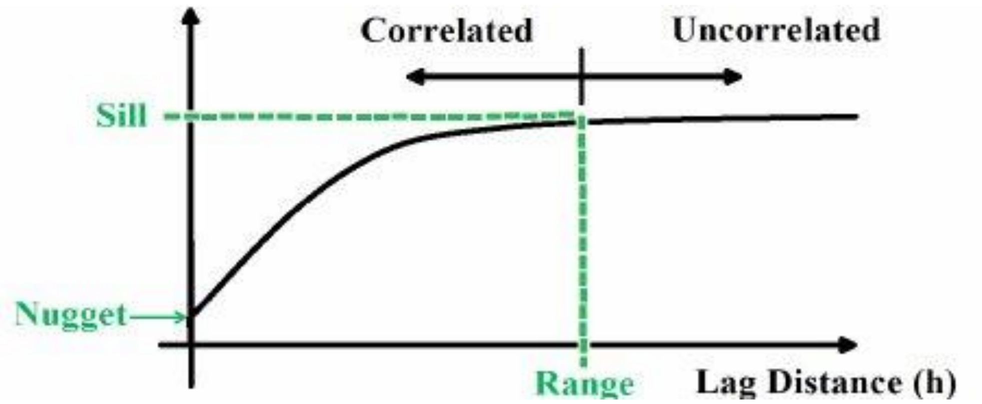



Figure 3.20: Typical semivariogram along with its components
(Baba et al., 2015).

### 3.3.3 Empirical analysis

Empirical analysis applied for examining the data set and applying statistical or analytical methods to draw conclusions or uncover patterns and relationships between the dependent and explanatory factors to show their relationship and contribution on the spatial variety in housing prices.

Empirical analysis where implemented on two stages:

- Applying OLS regression model.
- Applying GWR regression model.

### i Ordinary Least Squares (OLS)

In the conventional linear regression model, parameter estimation is commonly performed using the Ordinary Least Squares (OLS) method. OLS linear regression is utilized to predict or model a dependent variable based on its relationships with a set of explanatory variables (URL13). Conventional regression models do not explicitly consider spatial interactions and presuppose a constant process of price formation in geographical space. Consequently, the significance of parameters is independent of the spatial structure of the phenomenon being studied, potentially leading to a misinterpretation of the results. The standard Ordinary Least Squares (OLS) model used in this study is specified as follows in equation 3.2:

$$\mathrm{Y_i} = \beta_0 + \Sigma(\beta_k * \mathrm{Variable}_{ik}) + \varepsilon_i \qquad (3.2)$$

Where $Y_i$ represent the housing price for the i-th sale. In the regression equation, $\beta_0$ denotes the intercept (constant) term, $\beta_k$ represents the estimated parameter for each explanatory variable, where k ranges from 1 to 12. $\mathrm{Variable}_{ik}$ signifies the k-th variable for the i-th housing sale, and $\varepsilon_i$ is the error term for the i-th housing sale, capturing the unexplained variation in housing prices.

### a) Model performance

The Multiple R-Squared and Adjusted R-Squared, both from OLS, measure model performance by explaining data variation. Ranging from 0.0 to 1.0, Adjusted R-Squared is slightly lower as it accounts for model complexity, making it a more accurate indicator. Adding more variables tends to increase Multiple R-Squared but may decrease Adjusted R-Squared, emphasizing the need to select relevant variables to avoid overfitting.

### b) Stationarity

The Koenker (BP) Statistic, also known as Koenker's studentized Breusch-Pagan test, checks if the relationship between a model's explanatory and dependent variables is consistent across geographic locations (stationarity) and variable magnitudes (homoscedasticity) (Zhao et al., 2018). To address stationarity issues, Geographically Weighted Regression (GWR) can be used to ensure ifءلال residual variance is constant across predicted values. In the second OLS model, the residual vs. predicted scatterplot showed more randomness than the first, but GWR is still necessary to meet the heteroscedasticity assumption.

### c) Independency

The assumption of independence states that the observations in the dataset are independent of each other. In the context of OLS regression, this implies that the errors or residuals of the model are not correlated (Chaudhary et al., 2023). Violation of this assumption can lead to biased and inefficient coefficient estimates. This assumption coincide with the stationarity assumption, in order to test assess this assumption, The Spatial Autocorrelation (Moran's I) tool was employed on the regression residuals to verify their spatial randomness.

### d) Linearity

The OLS model is made on the assumption of linearity, indicating that alterations in the independent variables are linked to proportional changes in the dependent variable. Linearity can be tested by assessing the histogram for the explanatory variables distribution and their relationship with the dependent prices variable.

#### e) Normality

Ordinary Least Squares (OLS) assumes that residuals (the differences between observed and predicted values) follow a normal distribution (Schmidt & Finan, 2018). Testing for normality helps assess if OLS is suitable. If residuals significantly deviate from normality, alternative methods like robust regression or generalized linear models may be better.

In this study, R was used to fit distributions to the housing price data in order to assess its underlying distributional characteristics. To evaluate normality, tests such as Shapiro-Wilk and Anderson-Darling were applied.In the case of non-normal distribution, two modeling options could be implemented: the use of non-parametric regressions like quantile regression or locally weighted scatterplot smoothing (LOESS), or performing residual analysis to evaluate if the model is adequately capturing the data. If patterns or heteroscedasticity are found in the residuals, alternative methods should be considered.

The Jarque-Bera test, part of OLS diagnostics, also checks for normality in the model's errors. A low p-value (below 0.05) suggests non-normality, indicating possible bias in the model.

### ii Geographically Weighted Regression (GWR)

Geographically Weighted Regression known as one of the representing spatial regression methods that are usually applied in studying phenomena and topics related to geographical location through analyzing and evaluating a local model of the independent variable, which in the case of this study is housing prices, to understand or predict the existence of spatial variation in these prices and how they are affected by various factors. The GWR regression model can be described as an expansion of the conventional linear regression model achieved by incorporating spatial relationships by assigning weights to individual observations based on their locations. Stemming from non-parametric regression, its core involves the establishment of local linear regressions at each point where measurement data is present. This method fits a regression equation to each feature of the dataset. For each feature within a specific bandwidth, the information of the independent variable and the explanatory variables located within the specified bandwidth of this feature are combined to form this equation. The shape and extent of each bandwidth analyzed are based on the bandwidth type and bandwidth selection method parameters. The GWR model applied based on equation 3.3:

$$y_i = \beta 0(u_i, v_i) + \sum_{j=1}^{p} \beta_j (u_i, v_i) x_{ji} + \varepsilon_i \qquad (3.3)$$

where $y_i$ is the dependent (housing prices) variable at location i, $\beta 0(u_i,v_i)$ is the spatially varying intercept term, $\beta_j(u_i,v_i)$ is the spatially varying coefficient for the j-th explanatory variable at location i, $x_{ji}$ is the value of the j- Mth explanatory variable at location i, $\varepsilon_i$ is the error term at location i, and $(u_i,v_i)$ represents the spatial coordinates of location i.

First, The cross-validation (CV) technique as well as the Akaike Information Criterion (AIC) were utilized. in order for the model to determine the optimal bandwidth based on the nature of the dataset, which produced an error due to the local collinearity in the dataset (many records of the housing prices are actually for the same building but for different apartments within that building, and since all the near values were moved to the center of their nearest building it caused for these values to be located at the same point).

An outlier is a value that is different and lies at an isolated distance from other values in a given data set. Detecting and properly dealing with outliers is essential because outliers show significant changes in both the mean and data’s standard deviation. One of several techniques used to handle data that contain outliers is the interquartile range technique, “The interquartile range lies between the distribution’s first and third quartiles” (Mallikharjuna Rao et al., 2023) as described in equation 3.4.

$$\mathrm{IQR} = \mathrm{Q1} - \mathrm{Q3} \tag{3.4}$$

Where IQR is the interquartile range, Q1 is the first quartile ( the 25th percentile of the dataset) and Q3 is the third quartile ( the 75th percentile of the dataset). OLS3, OLS4 and GWR2 were applied to the final dataset after excluding outlier values. The explanatory variables distance to malls and distance to ATM were excluded this time when applied GWR2 because they where found to be insignificant for the model prediction in the applied OLS3 regression model.

# CHAPTER 4

# DEVELOPED GIS APPLICATION

## 4.1 Introduction to the GIS Application

To facilitate an in-depth spatial analysis of housing prices across Ankara, a customized Geographic Information System (GIS) application was developed. The primary objective of this application is to visualize the distribution of housing prices across different districts and neighborhoods, by providing a system that enables urban planners and policymakers to understand the spatial dynamics of housing markets and the impact of several amenities on housing prices. By incorporating various explanatory factors into the analysis, the application aims to provide insights into how these factors influence housing prices. This GIS tool also serves as a user-friendly platform for GIS beginners, allowing them to easily generate heat maps and graphs, thereby enhancing their ability to interpret spatial data.

## 4.2 Tools and Technologies

The GIS application was built using MapInfo and Visual Studio, leveraging their robust spatial analysis and development capabilities. The programming language chosen for this project was C#, owing to its compatibility with both MapInfo and Visual Studio, and its ability to handle complex spatial computations efficiently. The combination of these tools enabled the creation of a powerful and interactive application that meets the project's analytical and visualization requirements.

## 4.3 Concept of Zoning

A key feature of this application is the use of Thiessen polygons, also known as Voronoi polygons, which play an important role in spatial data analysis. Thiessen polygons are geometric constructs that partition a space into non-overlapping regions based on the proximity to a specific set of points. Each polygon contains one generating point, and every location within a polygon is closer to its generating point than to any other point. This method is highly effective in spatial analysis as it allows the creation of distinct areas influenced predominantly by their central points, providing a clear delineation of spatial influence zones.

## 4.4 Application of Zone Creation in Housing Price Analysis

In this study, Thiessen polygons were employed to generate spatial zones around averaged housing transaction points across Ankara. By averaging housing prices within each zone, the polygons provided a means to visualize how housing prices are distributed spatially. The polygons were used to compute average values for overlapping housing transactions, resulting in zone areas that accurately represent the spatial distribution of these transactions. This technique ensures that each zone reflects the influence of its central housing transaction, offering a precise visualization of the spatial dynamics in the housing market. Thiessen polygons can be merged based on housing price ranges, in this study this zoning technique simplifies the visualization and analysis of spatial price patterns in Ankara. By grouping adjacent zones with similar price characteristics, this method creates larger, more cohesive areas that represent broader market trends, making it easier for urban planners to interpret housing price dynamics as presented in Figure 4.1. This approach not only improves zone management but also enhances the ability to compare any region whether it is districts, neighbors, or any manually selected region and predict future trends, offering valuable insights for targeted development and housing policy interventions. Table 4.1 present the five prices ranges in Çankaya.

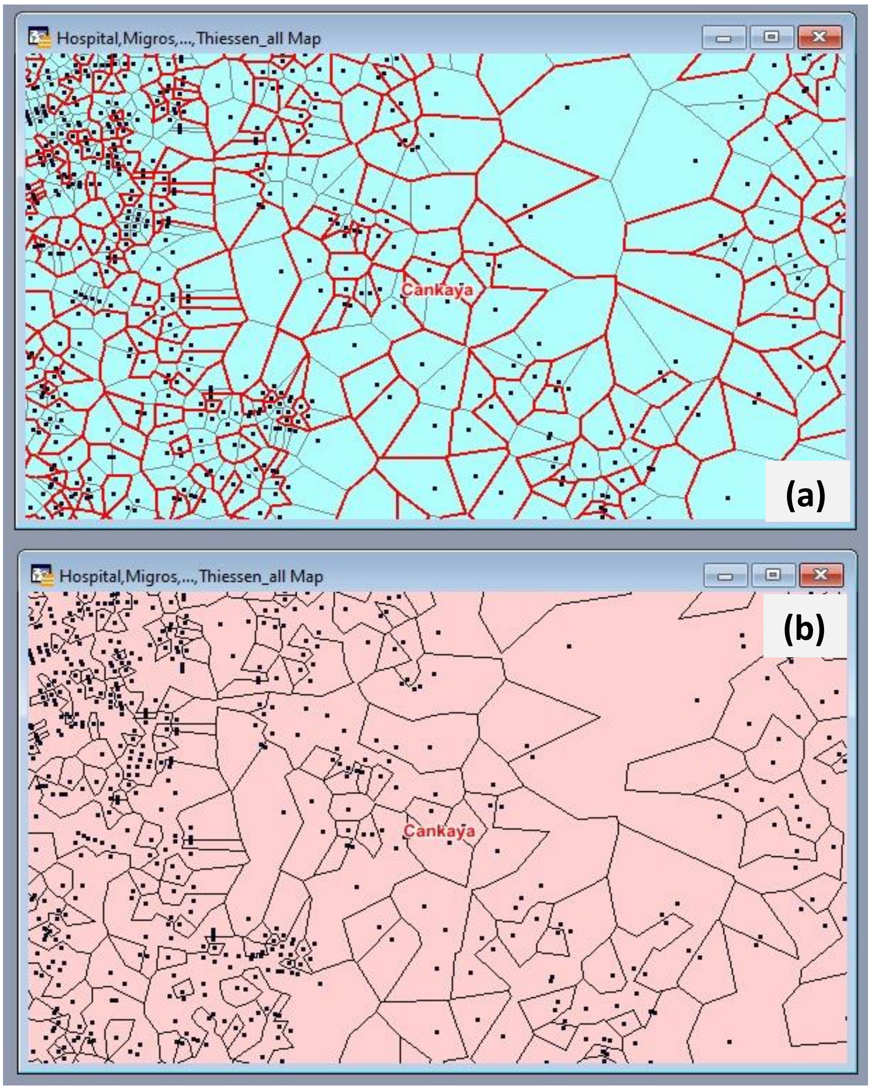



Figure 4.1: (a) Thiessen polygons representing housing transactions in central Çankaya (in gray) and merged Thiessen zones based on five housing price ranges in the same area (in red). (b) Merged Thiessen zones based on five housing price ranges in the central Çankaya.

Table 4.1: Housing price ranges in Çankaya.

| Price Range | Price Interval (USD) |
|---|---|
| PRange1 | 26,190 - 69,255 |
| PRange2 | 69,255 – 96,525 |
| PRange3 | 96,525 – 126,594 |
| PRange4 | 126,594 – 166,320 |
| PRange5 | 166,320 – 218,700 |

## 4.5 Classification of Housing Price Ranges

To further analyze the distribution of housing prices, the generated zone areas were categorized into five distinct price ranges (class intervals). These ranges were identified to help demonstrate how various explanatory factors (such as proximity to amenities) are distributed spatially within each price category. Experimentation with different numbers of intervals (5, 7, and 10) revealed that using five classes best captured the variations in housing prices and the influencing factors within the Ankara metropolitan area. This classification approach enabled the effective analysis and visualization of the relationship between housing prices and included explanatory factors in the application.

## 4.6 Visualization Techniques and Spatial Distribution

The application integrates various visualization techniques to represent the data effectively. Heat maps were generated to illustrate the concentration and distribution of housing prices, with higher intensities indicating higher prices. These heat maps provide a visual summary of how housing prices vary across different regions of Ankara, highlighting areas of high and low price concentrations.

Additionally, the Natural Breaks method was utilized to create thematic maps, which compared housing prices to population percentages. The Natural Breaks method, also known as Jenks optimization, identifies natural groupings inherent in the data. This method calculates class breaks where there are significant jumps in data values, maximizing the differences between classes while minimizing the differences within them. By doing so, it effectively segments the data into defined classes, highlighting significant variations and ensuring that each class represents a distinct range of values. This approach is particularly useful for thematic mapping, as it ensures that similar values are grouped together, making the visualization more intuitive and informative.

## 4.7 Application Features and User Interaction

The GIS application was developed with a focus on user interaction and ease of use. Users can select different price ranges and view the corresponding spatial distribution of explanatory factors across Ankara's districts. The application provides interactive features such as zooming, panning, and clicking on any region to reveal detailed information about each area. These features enable users to explore the data dynamically, gaining deeper insights into the housing market's spatial dynamics within different scales.

Furthermore, the application allows users to generate histograms for each price range, displaying the distribution of various explanatory factors. This functionality helps users understand how factors such as distance to amenities correlate with housing prices. By providing a comprehensive and interactive visualization platform, the application enhances users' ability to analyze and interpret spatial data effectively.

### 4.7.1 Graph Tool: Comparing Housing Prices With Explanatory Factors

One of the primary tools integrated into the GIS application is a graph tool that visualizes the relationship between housing prices and the explanatory factors identified in this study. This tool generates graphs that compare housing prices within each selected district against variables such distance to amenities, and other socio-economic factors such as population percentage between 2014 to 2021. By visualizing this data, users can easily identify patterns and trends in housing prices relative to these factors. For instance, Figure 4.2 illustrates the application interface, while Figures 4.3 and 4.4 demonstrate how this tool can be used to create heat maps showing the distribution of housing prices and population across Ankara. These visualizations help pinpoint areas with significant price changes and correlate them with changes in the explanatory factors.

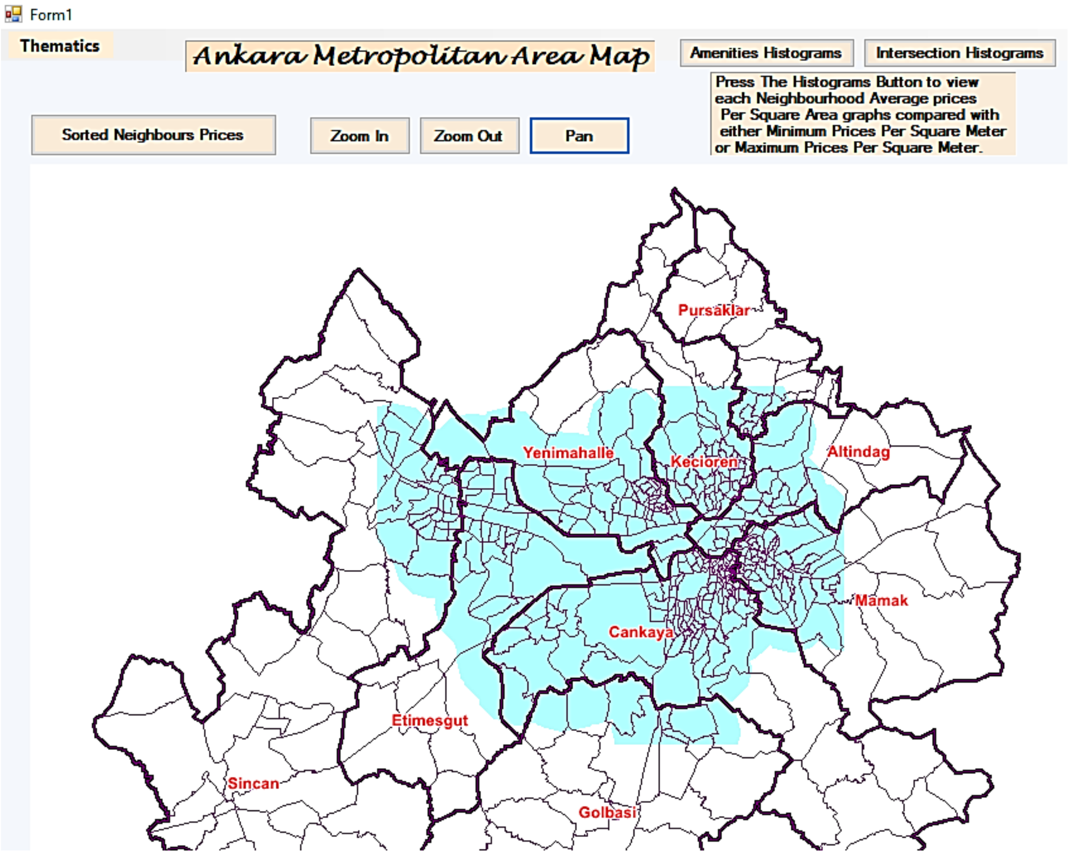



Figure 4.2: Ankara metropolitan map within neighborhood level.

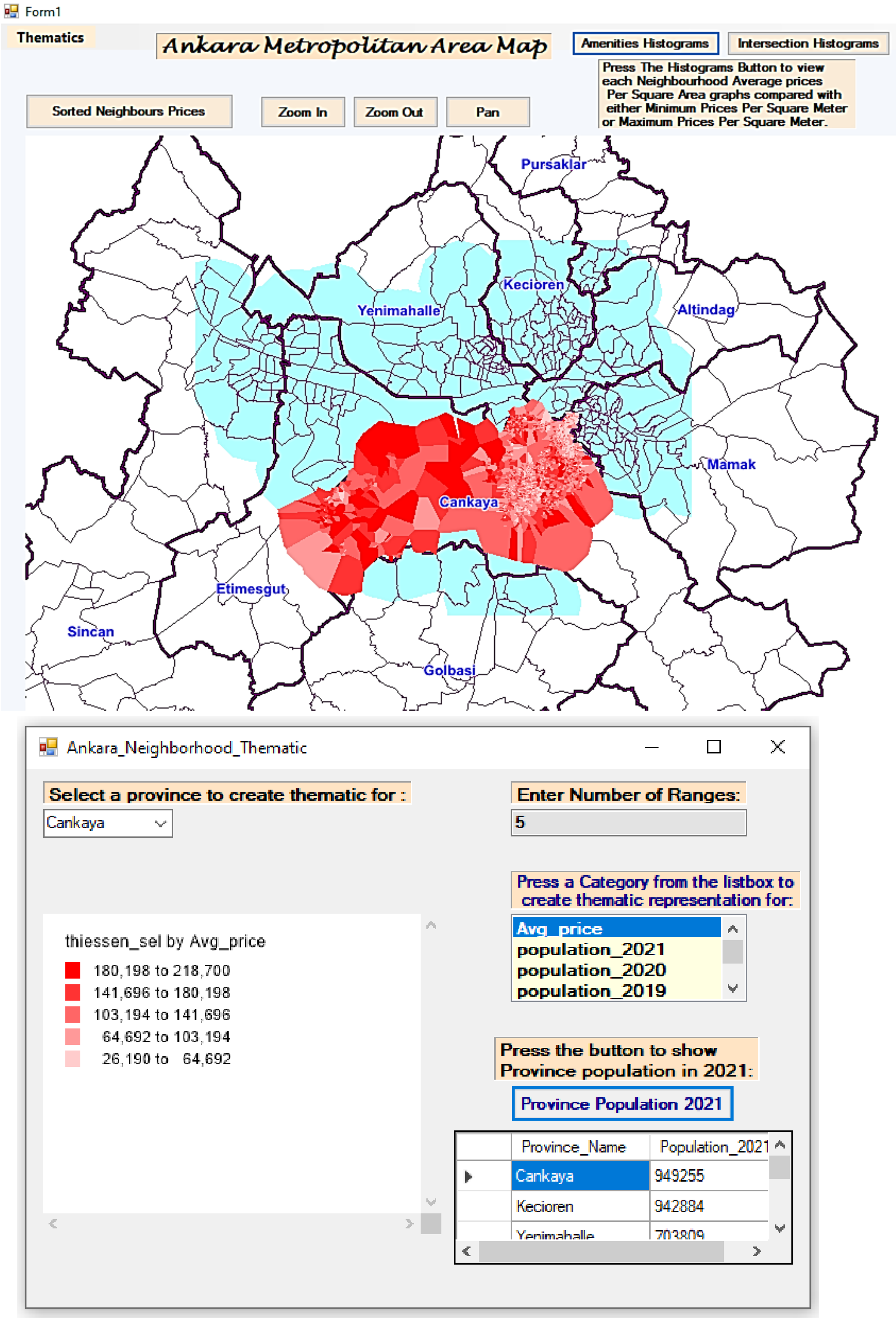



Figure 4.3: Housing prices heat map for Çankaya.

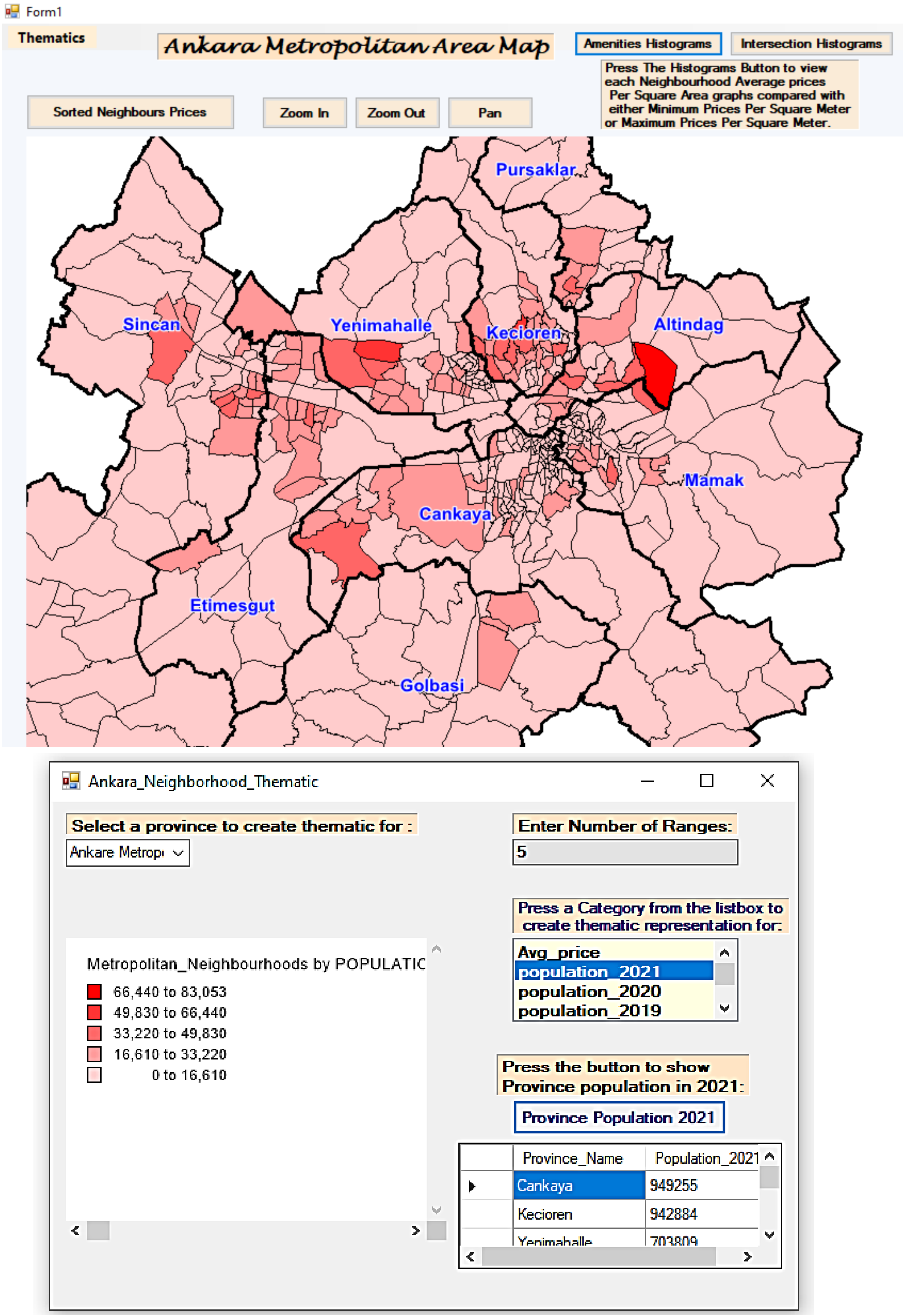



Figure 4.4: Population in 2021 heat map.

### 4.7.2 Histogram tool: Distribution of explanatory variables by price range

To facilitate the understanding of how dependent (housing prices) and explanatory variables relate to one another, a histogram tool was developed and integrated into the Visual Studio interface. This tool enables users to view the distribution of various explanatory variables within different price intervals, offering a graphical representation of how these factors vary across different price ranges. By displaying histograms of explanatory variables for each price interval, as shown in Figure 4.5, the tool simplifies further analysis by highlighting the impact of these factors on housing prices in different districts. This visualization aids in understanding the broader trends and relationships between housing prices and variables such as proximity to amenities and other socio-economic indicators that can be imported to the application.

### 4.7.3 Intersection analysis tool: Distribution of explanatory variables by neighborhood

Another feature of the GIS application is an intersection analysis tool, which allows users to analyze the distribution of explanatory variables across multiple districts by selecting several neighborhoods. This tool computes the total number of each explanatory variable within the intersection area of the selected neighborhoods and the zones resulted from Thiessen polygons. The results are displayed as histograms, as shown in Figure 4.6. This capability enables users to examine how factors such as the availability of amenities, public services, and other influential elements are distributed across different neighborhoods and on a smaller scale level for more detailed analysis. By doing so, it provides valuable insights into how these factors contribute to housing price variations in the Ankara metropolitan area and into smaller zoned areas.

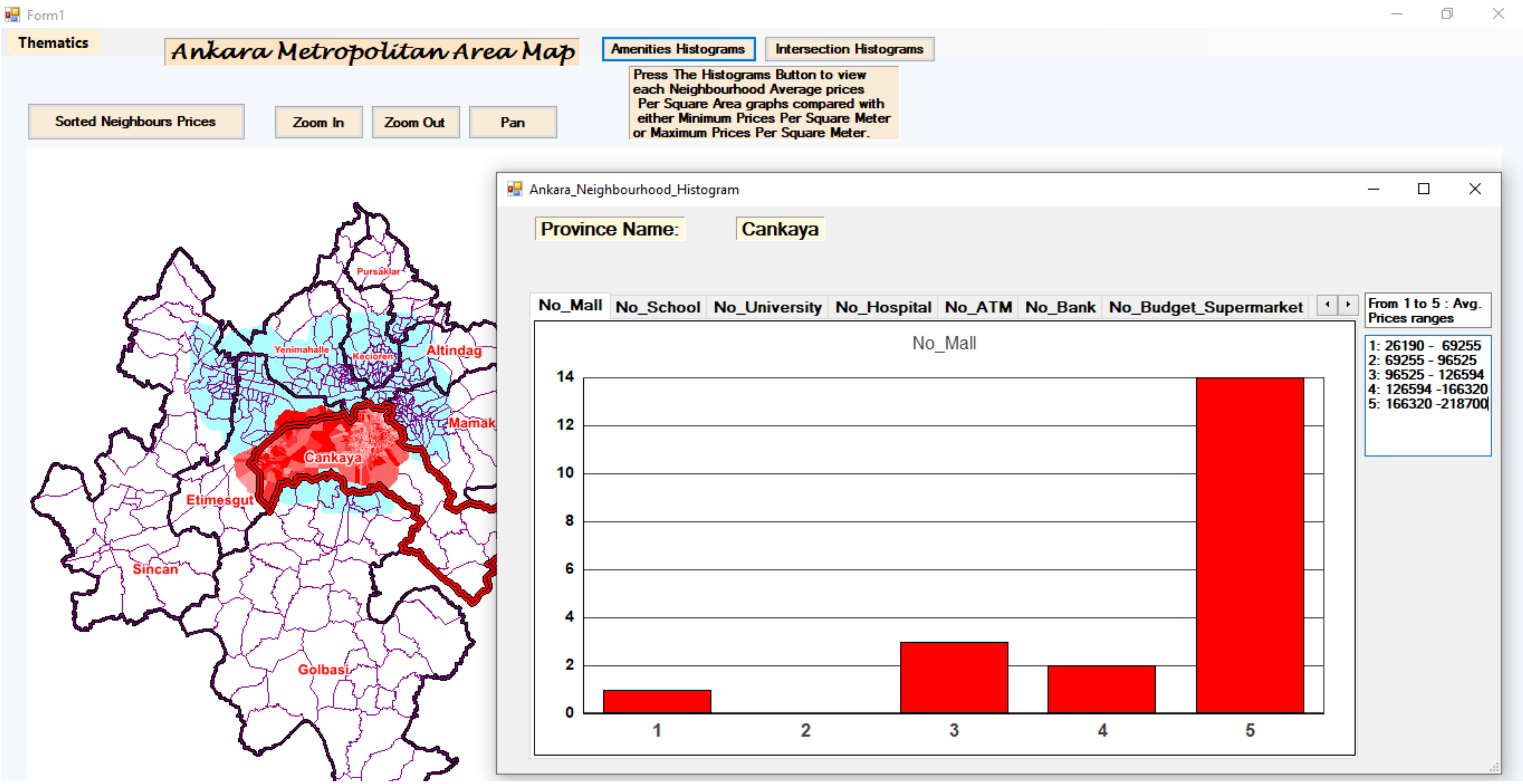



Figure 4.5: The histogram of total number of malls within Çankaya district in five housing prices ranges in USD.

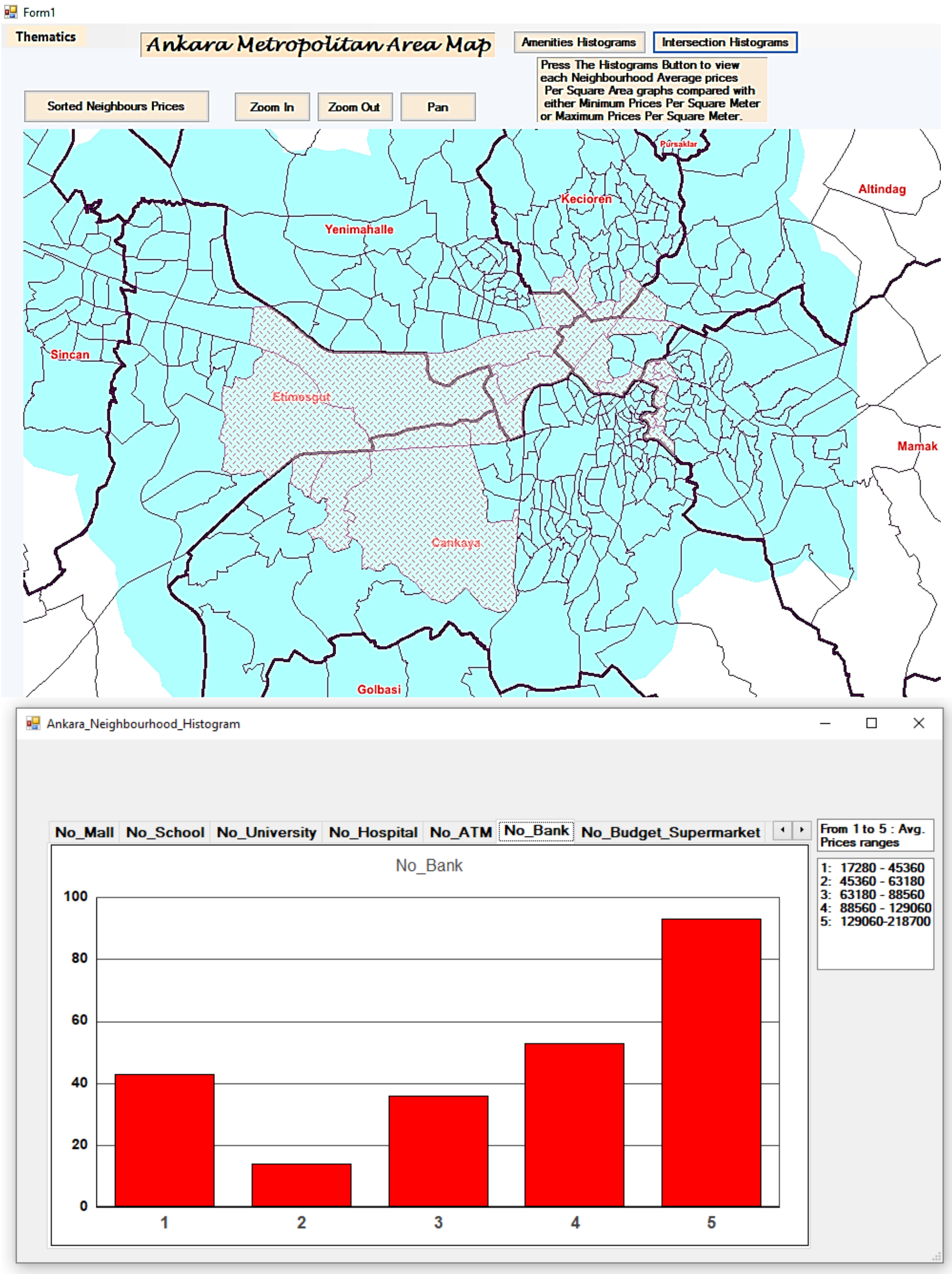



Figure 4.6: The histogram of total number of banks within several selected neighborhoods from several districts in five housing prices ranges in USD.

### 4.7.4 Sorting and zooming tool: Detailed price analysis by neighborhood

To facilitate detailed price analysis and navigation within the map, a sorting and zooming tool was added to the application interface. This tool allows users to sort the average and median prices per square meter from the highest to the lowest values. Additionally, users can select specific neighborhoods or multiple neighborhoods from a list and automatically zoom in on these areas within the map. Figure 4.7 illustrates the interface of this tool, showcasing how it enhances user interaction with the spatial data. This functionality is particularly useful for urban planners and policymakers who need to focus on specific areas for targeted analysis and decision-making.

### 4.7.5 Symbology tool: Visualization of explanatory factors on thematic maps

To enhance the interpretability of the spatial data, the GIS application incorporates a symbology tool. This tool uses appropriate symbols to represent each explanatory factor on top of the thematic map of housing prices for any selected district. This layered visualization method helps to illustrate the relationship between housing prices and factors such as distance to supermarkets, parks, schools, and other amenities. For example, Figure 4.8 demonstrates that budget supermarkets are more prevalent in lower-priced areas of Çankaya but are scarce in high-priced regions, suggesting a correlation between the availability of budget amenities and housing prices. This tool provides users with a clear visual representation of how various factors influence housing prices across different parts of Ankara.

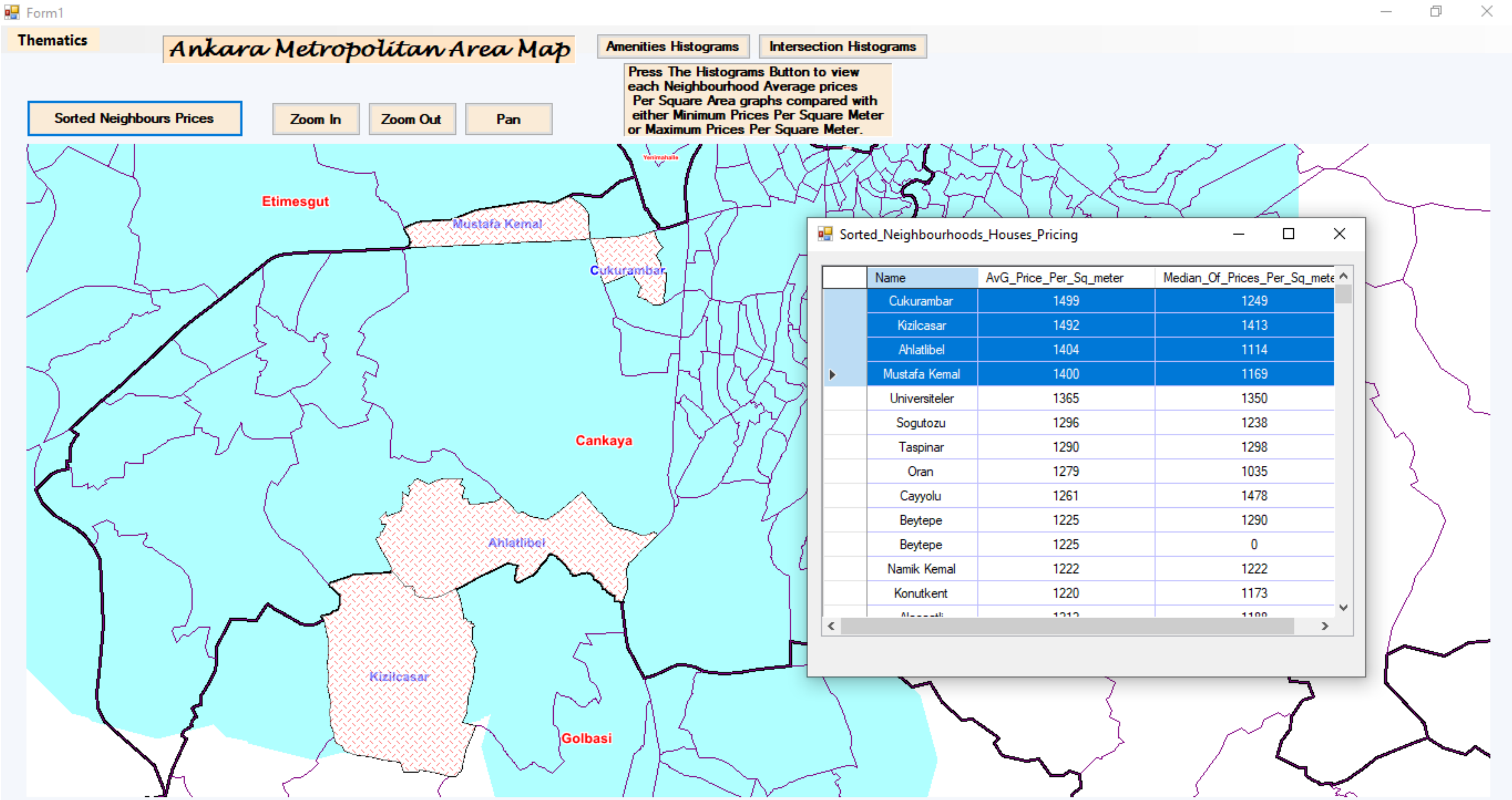



Figure 4.7: Sorted average prices per square meter, highliting  the four highest prices for neighbourhoods.

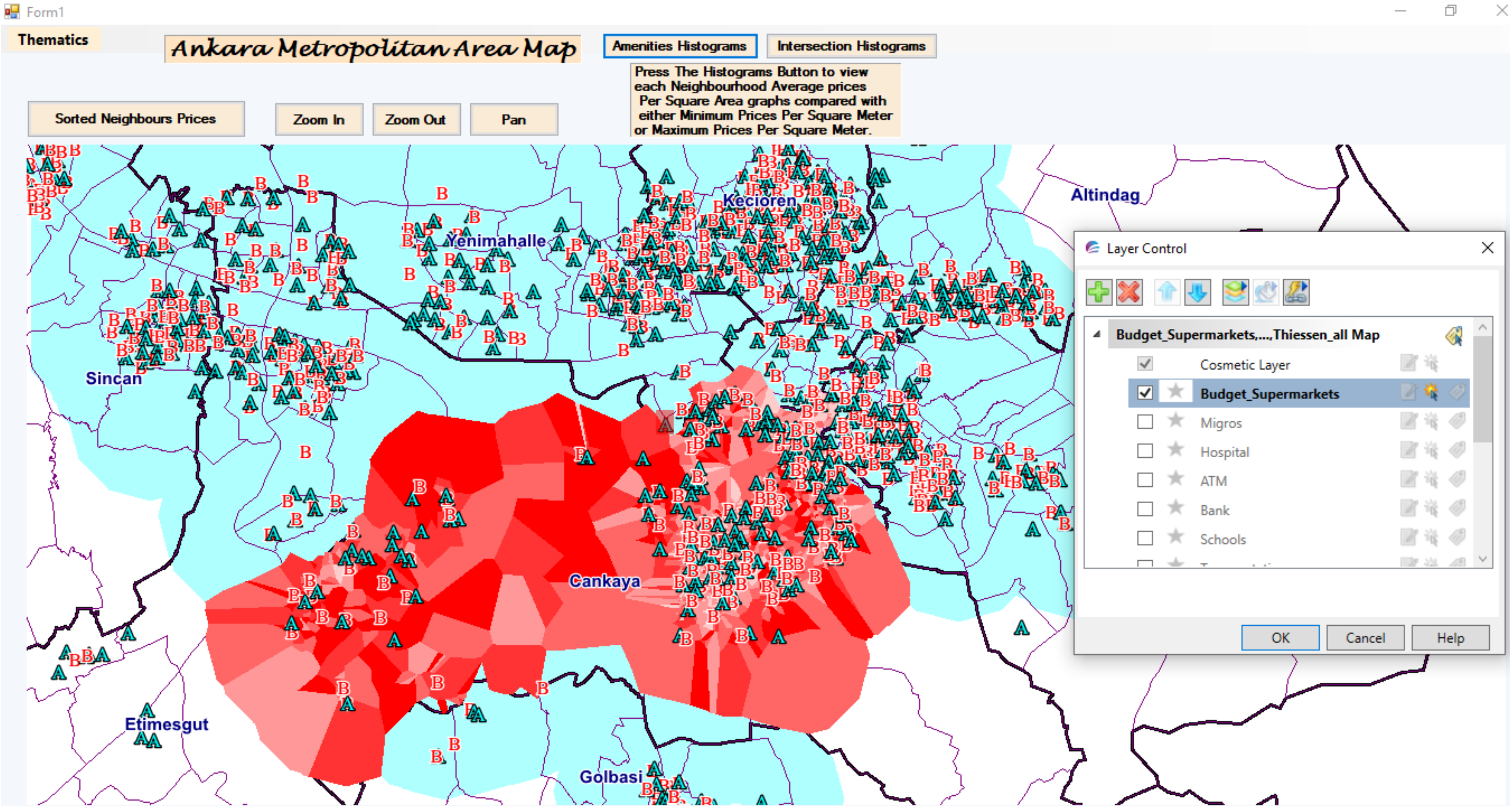



Figure 4.8: Distribution of budget supermarkets (BIM & A101) in Ankara metropolitan area.

### 4.7.6 Heat map analysis: Identifying high-priced regions

The GIS application also includes a heat map analysis tool, which visually represents the concentration of housing prices across different regions. The heat maps generated using this tool reveal that the western region of the Ankara metropolitan area contains the highest-priced houses. This visual tool helps users understand the spatial distribution of housing prices at a glance, highlighting areas of high and low price concentration. Analyzing the distribution of explanatory variables in these high-priced regions, it was found that proximity to hospitals and universities is a significant factor. For example, the western side of Ankara shows shorter distances to these amenities, contributing to higher housing prices in these areas. This insight helps explain the spatial patterns of the housing market and guides future urban planning and policy decisions.

### 4.7.7 Review of application code

The development of these tools and their integration into the GIS application involved complex coding using C#. The complete application code, including the implementation of the various tools and features described above, can be reviewed in Appendix C. This appendix provides detailed documentation of the code structure, functions, and algorithms used, offering insights into the technical aspects of the application development.

# CHAPTER 5

# RESULTS

## 5.1 Clustering Analysis Results: Morans'I and Krigging

Applying the Kriging geostatistical technique as shown in Figure 5.1 showed that the model starts to level out at a distance of approximately $1.724*10^4$ meter, almost 17 kilometer, which means the methodology is almost considering most of the pricing data as correlated points on the prediction procedure since that the diameter of the centered and most dense pricing data available in the Ankara metropolitan area is approximately 22 kilometer as illustrated in Figure 5.2.

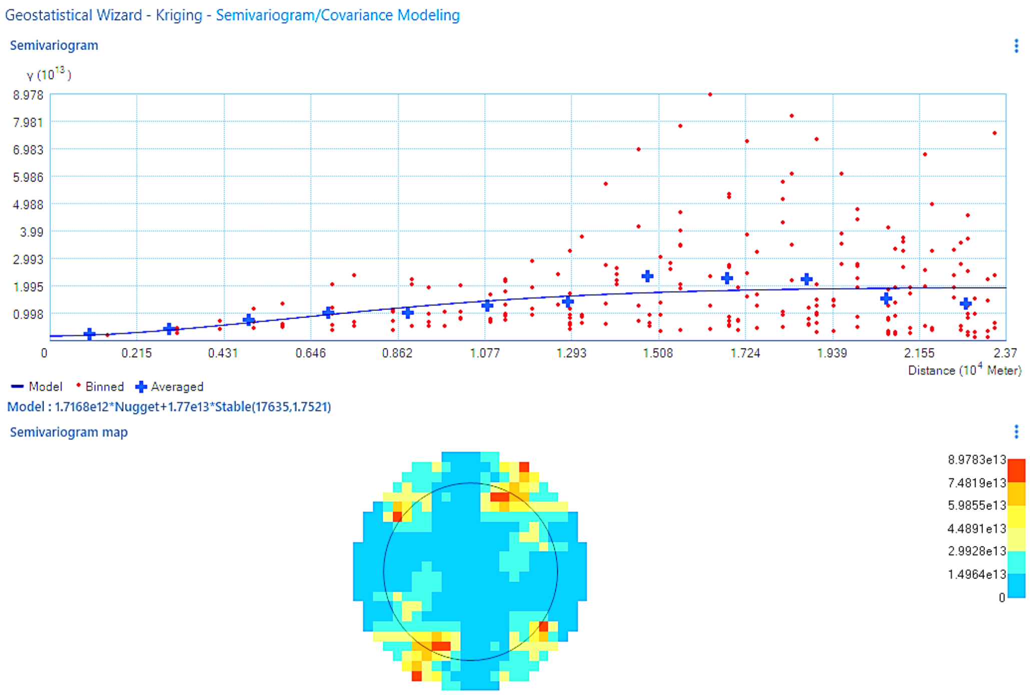



Figure 5.1: Semivariogram model for applying ordinary kriging on housing prices.

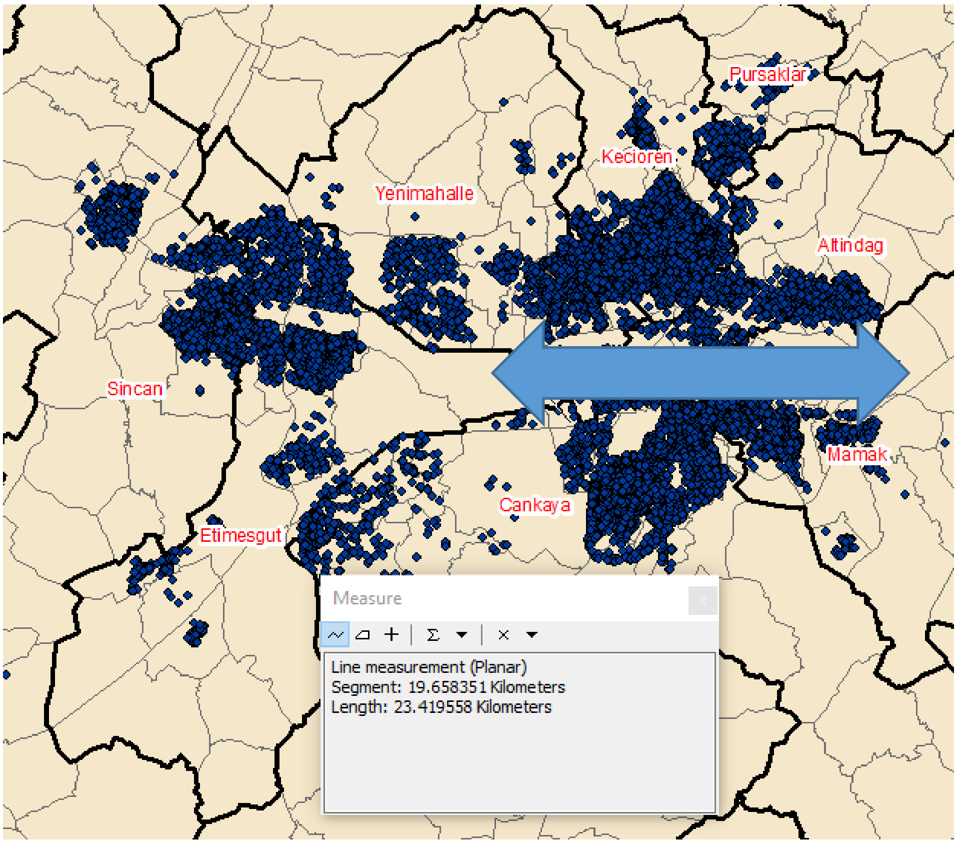



Figure 5.2: Diameter of Ankara metropolitan area.

In this study prices histogram showed clearly that there are very high values located far away from the mean of the prices, which makes the Kriging useful in this case. Therefore, under the assumption that a neighborhood diameter can range from 1 to 5 km, this study applied buffer zones of 1 km, 2 km, 3 km, 4 km, and 5 km. Each record within these buffer zones was compared with all neighboring records within the respective zones.

Table 5.1 shows that applying 1 km, 2km, 3km, 4km, and 5km as a buffer zones in the morans'I analysis showed that the p-value was < 0.01 and the z-score within the rang < -2.58 or > +2.58, which indicates that the housing transactions analyzed in this study exhibit significant clustering characteristics and positive spatial autocorrelation and are not a result of a random patterns the figures resulted from applying morans'I are illustrated in the Appendix D subsection.

Table 5.1**:** Global Moran I for Prices.

| | 1 km | 2 km | 3 km | 4 km | 5 km |
|---|---|---|---|---|---|
| **Moran's Index:** | 0.912 | 0.866 | 0.834 | 0.810 | 0.79 |
| **Expected Index:** | -0.000024 | -0.000024 | -0.000024 | -0.000024 | -0.000024 |
| **Variance:** | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| **p-value:** | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |

## 5.2 Empirical Anlalysis Results

In this study, four Ordinary Least Squares (OLS) models (OLS1, OLS2, OLS3, OLS4) and two Geographically Weighted Regression (GWR) models (GWR1, GWR2) were applied to the entire dataset to examine the impact of various explanatory factors on housing prices. The analysis was then repeated on separate datasets for the Çankaya district and the rest of the districts, with two GWR models applied to each.

For the whole dataset, OLS1 identified Area/m² and distance to ATMs as non-significant variables, leading to their exclusion in OLS2 and GWR1. After removing outliers, OLS3 further identified distance to malls and distance to ATMs as non-significant, which were then excluded in OLS4 and GWR2. Given the non-stationarity of the data, OLS was not applied to the separated datasets, and the GWR models excluded the previously identified non-significant variables.

As shown in Table 5.2, a total of four OLS models (OLS1, OLS2, OLS3, OLS4) and two GWR models (GWR1, GWR2) were applied to the entire dataset. Additionally, when the dataset was separated into Çankaya district and the rest of the districts, two GWR (GWR3 & GWR4) models were applied to each subset.

Table 5.2: Overview of Applied OLS and GWR Models.

| **Model** | **Dataset** | **Variables Excluded** | **Purpose** |
|---|---|---|---|
| | | | |
| **OLS1** | Study Area- Raw Data | --- | All variable model |
| **OLS2** | Study Area- Raw Data | Area/m², Distance to ATMs | Refined model w/o non-significant variables |
| **OLS3** | Study Area- Cleaned-up Data set* | --- | All variable model |
| **OLS4** | Study Area- Cleaned-up Data set* | Distance to malls, Distance to ATMs | Refined analysis after excluding additional non-significant variables |
| | | | |
| **GWR1** | Study Area- Raw Data | Area/m², Distance to ATMs | Spatial analysis before excluding outliers |
| **GWR2** | Study Area- Cleaned-up Data set* | Distance to malls, Distance to ATMs | Spatial analysis after excluding outliers |
| **GWR3_ Çankaya** | Cleaned-up Data set*- Çankaya | Area/m², Distance to malls, Distance to ATMs | Spatial analysis focused on Çankaya district |
| **GWR4_no Çankaya** | Cleaned-up Data set*- Study Area w/o Çankaya | Area/m², Distance to malls, Distance to ATMs | Spatial analysis focused on the rest of the districts |
| **All Variables: Area / (m2) ,Room Number , Floor Number, Dis. to University (m), Dis. to Hospital (m), Dis. to Mall (m), Dis. to Transport (m), Dis. to School (m), Dis. to ATM (m), Dis. to Bank (m), Dis. to BIM &A101 (m), Dis. to Migros (m)** | | | |

### 5.2.1 OLS1 & OLS2 results

**Model performance**

In the applied OLS regression model to analyze housing prices, where the number of house transactions in each neighborhood is the dependent variable, the model achieved an Adjusted R-Squared value of 0.42, which indicates that approximately 42% of the variability in housing transactions can be explained by the model of the explanatory variables. In other words, the model provides insights into about 42% of the factors contributing to housing prices. By assessing the explanatory variables, the resulted report indicated that ten out of twelve of the explanatory variables are significant and therefore important to apply in the model for better explanation of the housing prices transactions. Area per m2 and distance from ATM are the non-significant variables since that the p-value associated with their coefficients are >0.05. The Variance Inflation Factor (VIF) quantifies redundancy among explanatory variables. As a general guideline, explanatory variables with VIF values exceeding approximately 7.5 should be eliminated. The results indicated values that are not larger than 2.9 for the (VIF), suggesting the absence of multicollinearity among the explanatory variables which metioned in the summary OLS results section in Appendix E . In the OLS Diagnostics section, the Joint F-Statistic [e], Joint Wald Statistic [e], and Koenker (BP) Statistic [f] (Koenker's studentized Breusch-Pagan statistic) tests, which assess whether the explanatory variables in the model exhibit a consistent relationship with the dependent variable in both geographic and data space, are deemed statistically significant, as their associated p-values are <0.05.

The histogram of standardized residuals is displayed in Appendix E, offering a visual assessment of the residuals' distribution. This histogram is useful for detecting any deviations from normality, which could indicate potential issues with the model's assumptions.

A plot of OLS1 residuals versus predicted values is presented in Appendix E. This plot helps in identifying patterns or systematic structures in the residuals, which could suggest model misspecification or heteroscedasticity. A well-fitted model should display residuals scattered randomly around the horizontal axis, with no discernible pattern. The resulted plot for the Residual vs. Predicted from the OLS1 model revealed that the scatterplot has some structure towards the left side, which indicates that the model is performing well in areas characterized by low rates of housing transactions, but not doing well in locations with high rates of housing transactions.

**<u>Stationarity</u>**

The OLS model was applied twice. In the first application of OLS, the results indicated that the Koenker (BP) test was statistically significant. A review of the robust coefficient standard errors and the probabilities of the explanatory variables revealed that all variables, except for the area per square meter and distance from ATM, were effective and significant. In the second OLS model applied; the variables area per m2 and distance from ATM were eliminated, and all the remaining variables were statistically significant. The resulted plot for the Residual vs. Predicted from the first OLS model revealed that the scatterplot has some structure or pattern, which indicates that the assumptions of the Ordinary Least Squares (OLS) regression may not be fully satisfied. A descreption of the summary results OLS2 excluding the insufficiont varibles from OLS1 model presented in Appendix E.

The histogram of the standardized residuals obtained from OLS1 shows a distribution that is centered around 0 but shifted to the left. This indicates that the residuals exhibit negative skewness, meaning that there are more residuals with negative values than positive ones. This asymmetry suggests that the model may be underestimating the actual values for certain observations. Such a pattern could indicate issues with the model's assumptions or the presence of outliers or influential points that affect the overall fit. The scatter plot of OLS2 residuals versus predicted values indicates that the residuals are more evenly scattered around the center, compared to the OLS1 model. Unlike the OLS1 model, where the residuals exhibited a noticeable shift to the left, the OLS2 residuals do not show this skewness. This more balanced dispersion around the center suggests that the OLS2 model may provide a better fit to the data, with fewer systematic biases and a more symmetrical distribution of errors. This improvement is due to better specification of the model and removal of the irrelevant variables found in OLS1 . Both figures of residuals histograms and residuals versus predicted values can be reached out in Appendix E.

**Independency**

By applying Moran's I tool to the regression residuals to verify their spatial randomness, the results showed spatial autocorrelation in residuals which within the framework of regression analysis implies that the assumption of independence is violated, the results of morans'I can be viewed in Appendix E. The existence of spatial autocorrelation suggests that the residuals are in fact not truly independent, and there is some spatial pattern or structure that remains unaccounted for in the model.

**Linearity**

By assessing the histogram for the explanatory variables distribution and their relationship with the dependent prices variable, the histograms indicated that a positive linear relationship with each of the explanatory variables and the dependent variable are existed but for the Area per m2 variable from the OLS1 model which was found to be one of the insufficient variables for explaining the model the relationship is not linear as presented in Appendix E along with the histogram for the explanatory variables distribution and their relationship with the dependent prices variable after excluding the insufficient variables in OLS2.

**Normality**

Applying both Anderson-Darling and Shapiro-Wilk tests indicated a very small P-values, way smaller than 0.05 as mentioned in Table 5.3 which means that the null hypothesis that average price per square meter data comes from a normal distribution can be rejected with a high degree of confidence. Therefore, it can be concluded that the housing prices data is unlikely to follow a normal distribution. Despite this, Ordinary Least Squares (OLS) was applied because it is robust to deviations from normality, especially in large samples. Additionally, OLS provides useful baseline estimates and is commonly used for initial analysis even when normality is not strictly met.

Table 5.3: Testing normality results.

| **Distributions** | **AIC** | **BIC** | **likelihood** | **P-value Anderson-Darling test** | **P-value Shapiro-Wilk test** |
|---|---|---|---|---|---|
| **LogNormal** | 4212.07 | 4220 | -2104.036 | 3.70E-24 | 1.24E-15 |
| **Normal** | 4264.94 | 4272.86 | -2130.468 | 3.70E-24 | 1.24E-15 |
| **Exponantional** | 4526.62 | 4530.59 | -2262.311 | 2.35E-12 | 1.24E-15 |
| **Gamma** | 4184.18 | 4192.11 | -2090.089 | 2.08E-12 | 1.24E-15 |

The resulted p-value associated with the Jarque-Bera statistic [g] test (a test used to assess the normality of residuals in a regression model) is reported as 0.00000*, it indicates an extremely small p-value and the asterisk (*) denotes statistical significance. The Jarque-Bera test suggests that the assumption of normality for the residuals is violated. A p-value of 0.00000* indicates strong evidence against the null hypothesis that the residuals follow a normal distribution. Therefore, with such low p-value, it is concluded that the residuals in the OLS regression model do not conform to a normal distribution. This implies that there may be non-normal patterns or outliers present in the residuals, which may affect the reliability of the model's assumptions and the validity of the statistical inference. Therefore, considering GWR as an alternative modeling technique for investigating the nature of the residuals may help to address the violation of the normality assumption. Maps of the standard deviation of residuals distribution for the OLS1 and OLS2 models are presented in Appendix E**.**

## 5.3 GWR Results

To resolve the stationarity issue from applying OLS; Geographically Weighted Regression (GWR) can be applied given that the variance of the errors or residuals should remain constant across all levels of the independent variables. In other words, the dispersion of the residuals should be consistent across the entire range of predicted values. In the second applied OLS model the scatterplot of the Residual vs. Predicted where more randomly distributed than in the first applied OLS model, yet applying GWR is important to insure the heteroscedasticity assumption is not violated.

OLS assumes that the relationships between variables are constant across the study area and that residuals are homoscedastic meaning they have constant variance. However, these assumptions may not hold true in spatial data where relationships can vary by location, leading to spatial heterogeneity and non-constant residual variance. This is where GWR becomes necessary, as it allows for local variations in the relationships between variables, providing a more accurate and clear understanding of the data. GWR accounts for spatial non-stationarity by calibrating a regression model at each location, thereby addressing the limitations of OLS and improving model reliability and interpretation. The variables applied in the GWR model in which the area/$m^2$ and distance to ATM were excluded as in OLS2 are presented in Appendix F.

The results from applying the Geographically Weighted Regression (GWR1) model using 100 neighbors can be accessed in Appendix F. The model demonstrates a strong fit, with an $R^2$ value of 0.91 and an adjusted $R^2$ of 0.86, indicating that a significant proportion of the variance in the dependent variable is explained by the model. Additionally, the corrected Akaike Information Criterion (AICc) value of 301.85 suggests a good balance between model fit and complexity. The table also details the variables utilized in the GWR1 model, providing a comprehensive overview of the factors contributing to the spatial variation in the analysis.

### 5.4 OLS and GWR Excluding Outliers Results

After identifying and excluding outliers from the dataset, the Ordinary Least Squares (OLS) and Geographically Weighted Regression (GWR) models were reapplied to assess the impact of outlier removal on the robustness and reliability of the results. Outliers can significantly distort statistical analyses by skewing the results and affecting the accuracy of model estimations. By excluding these anomalous data points, the models can provide a clearer, more accurate representation of the underlying patterns in the data.

The revised OLS model, post-outlier exclusion, demonstrated improved statistical properties. The scatterplot of residuals versus predicted values showed a more random distribution, indicating a better fit of the model to the data. The Koenker (BP) test was re-evaluated and remained statistically significant, suggesting persistent issues with heteroscedasticity. However, the removal of outliers resulted in a reduction of variance in residuals, leading to more reliable coefficient estimates and more accurate p-values for the explanatory variables. This enhanced the validity of the OLS model in capturing the primary trends in housing prices.

Despite these improvements, the limitations of the OLS model in handling spatial heterogeneity and non-stationarity were still evident. To address these limitations, the GWR model was reapplied to the outlier-free dataset. GWR offers the advantage of accounting for local variations in relationships between the dependent and independent variables. By calibrating a separate regression equation for each location, GWR can reveal spatially varying relationships that OLS may overlook.

The application of GWR to the cleaned dataset provided deeper insights into the spatial dynamics of housing prices. The local coefficients from the GWR model highlighted significant geographic variations in the impact of explanatory variables, which were not apparent in the global OLS model. This underlined the importance of considering spatial heterogeneity in the analysis. By excluding outliers, the GWR model's ability to accurately reflect local variations was further enhanced, leading to more precise and context-specific interpretations of the data.

The histogram of housing prices excluding the outliers is illustrated in Figure 5.3. Histograms of the explanatory factors used in the spatial analysis, following the removal of outliers from the dataset are represented in the Appendix G. These histograms provides a visual representation of the distribution of each factor, illustrating how the data has been adjusted to exclude extreme values.

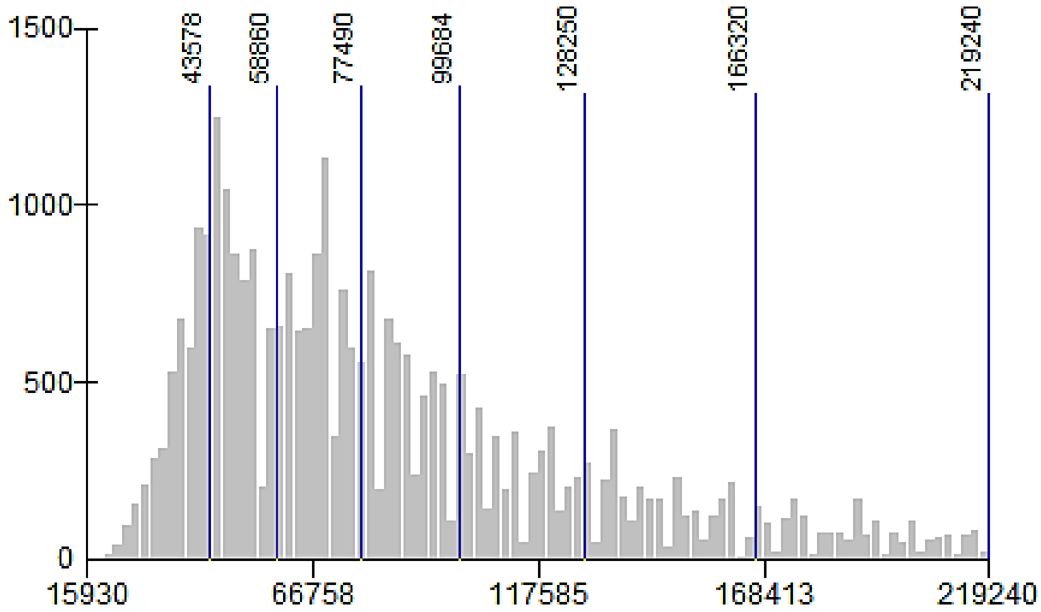



Figure 5.3: Histogram of housing prices in (USD) excluding outliers.

Similarly the statistical summaries for the explanatory factors were obtained after excluding the outliers from the dataset as described in Table 5.4.

Table 5.4: Statistical Summaries after excluding outliers (N= 30842).

| Variable | Mean | Std. Dev. | Min | Q1 | Median | Q3 | Max |
|---|---|---|---|---|---|---|---|
| Price | 83421.58 | 42304.61 | 15930 | 50166 | 72468 | 105300 | 219240 |
| Area Per m2 | 135.54 | 32.36 | 53 | 120 | 130 | 150 | 232 |
| Rooms Number | 3.96 | 0.72 | 1 | 4 | 4 | 4 | 10 |
| Floor Number | 2.89 | 3.75 | -4 | 1 | 2 | 4 | 30 |
| dist.to Transport | 141.10 | 79.97 | 8.16 | 79.97 | 127.17 | 187.50 | 387.55 |
| dist.to University | 4517.85 | 3184.72 | 13.75 | 1946.23 | 3768.66 | 6704.57 | 15595.64 |
| dist.to Mall | 2586.60 | 1535.47 | 7.37 | 1512.66 | 2285.10 | 3339.44 | 7079.58 |
| dist.to Hospital | 1573.19 | 1029.62 | 4.49 | 797.62 | 1358.07 | 2113.34 | 5069.02 |
| dist.to School | 251.17 | 165.50 | 0.68 | 140.78 | 220.76 | 318.80 | 1756.60 |
| dist.to ATM | 332.56 | 235.51 | 0.49 | 160.11 | 287.84 | 447.86 | 2454.47 |
| dist.to Bank | 791.39 | 561.96 | 1.93 | 379.75 | 675.49 | 1067.66 | 3871.22 |
| dist.to Migros | 1670.48 | 1234.66 | 1.15 | 663.78 | 1351.60 | 2443.40 | 6893.42 |
| dist.to Budget Supermarket | 226.74 | 207.85 | 0.35 | 106.62 | 178.03 | 280.97 | 3579.27 |

### 5.4.1 OLS3 results

After removing the outliers, the OLS3 model was applied, revealing that the factors 'distance to malls' and 'distance to ATM' were insufficient. Summary results for the OLS3 model variables, the histograms of the OLS3 explanatory variables and their relationship to the dependent variable, price, the histogram of standardized residuals for OLS3, the residuals versus predicted plot and the map of the standard deviation of residuals distribution for the OLS3 model are presented in Appendix H**.** The residuals looks distributed around the zero better than OLS1 and OLS2 models that contained the outliers.

### 5.4.2 Final OLS Model (OLS4) results

Subsequently, 'distance to malls' and 'distance to ATM' were excluded in the OLS4 model. The summary results for the OLS4 model variables are shown in Table 5.5. Figure 5.4 provides the histograms of the OLS4 explanatory variables and their relationship to the dependent variable, price. The histogram of standardized residuals for OLS4 is depicted in Figure 5.5, and Figure 5.6 presents the residuals versus predicted plot for OLS4. This progression highlights the iterative refinement of the model by excluding insufficient variables to improve the overall fit and accuracy.

Table 5.5: OLS4 summery results- Model variables.

| Variable | Coefficient [a] | StdError | Probability [b] | Robust SE | Robust t | Robust Pr[b] | VIF (c) |
|---|---|---|---|---|---|---|---|
| **Intercept** | -233139.7967 | 18916.901982 | 0.000000* | 22335.793991 | -10.437945 | 0.000000* | - |
| **Area/ (m2)** | 12678.396807 | 159.967831 | 0.000000* | 200.287681 | 63.300932 | 0.000000* | 3.082909 |
| **Room Number** | -56883.01264 | 7045.550145 | 0.000000* | 8587.994785 | -6.623550 | 0.000000* | 2.828103 |
| **Floor Number** | 62656.640560 | 869.191720 | 0.000000* | 1076.050012 | 58.228372 | 0.000000* | 1.177127 |
| **Near_Dist. to University** | -14.513956 | 1.089638 | 0.000000* | 1.035947 | -14.010322 | 0.000000* | 1.452031 |
| **Near_Dist. to Hospital** | -21.369840 | 3.006802 | 0.000000 | 3.110801 | -6.869563 | 0.000000* | 1.198203 |
| **Near_Dist. to Transport** | 244.491926 | 37.982267 | 0.000000* | 38.760337 | 6.307786 | 0.000000* | 1.084791 |
| **Near_Dist. to School** | 295.965623 | 19.556213 | 0.000000* | 22.567882 | 13.114462 | 0.000000* | 1.196676 |
| **Near_Dist. to Bank** | 135.261866 | 6.090104 | 0.000000* | 7.091430 | 19.073990 | 0.000000* | 1.352248 |
| **Near_Dist. to Migros** | -115.319849 | 2.789407 | 0.000000* | 2.773310 | -41.582033 | 0.000000* | 1.430622 |
| **Near_Dist. to Budget Supermarket** | 450.671852 | 17.635595 | 0.000000* | 22.558590 | 19.977838 | 0.000000* | 1.345679 |

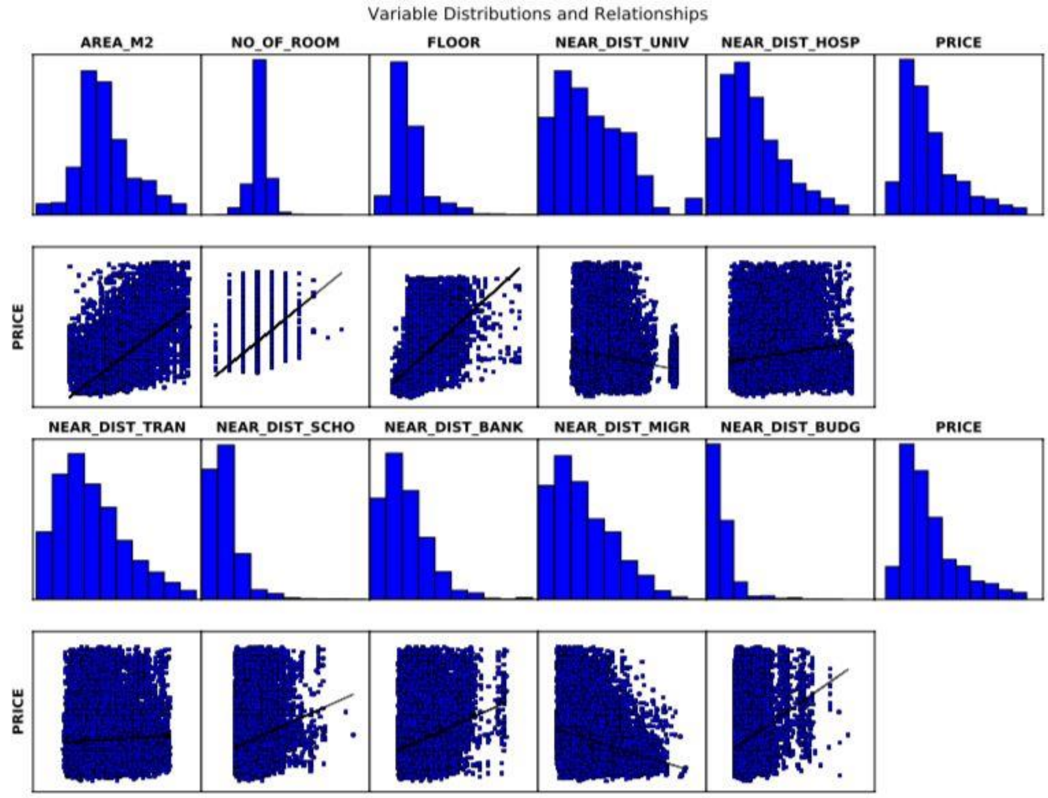



Figure 5.4: OLS4 explanatory variables histogram and their relationship to the dependent variable price.

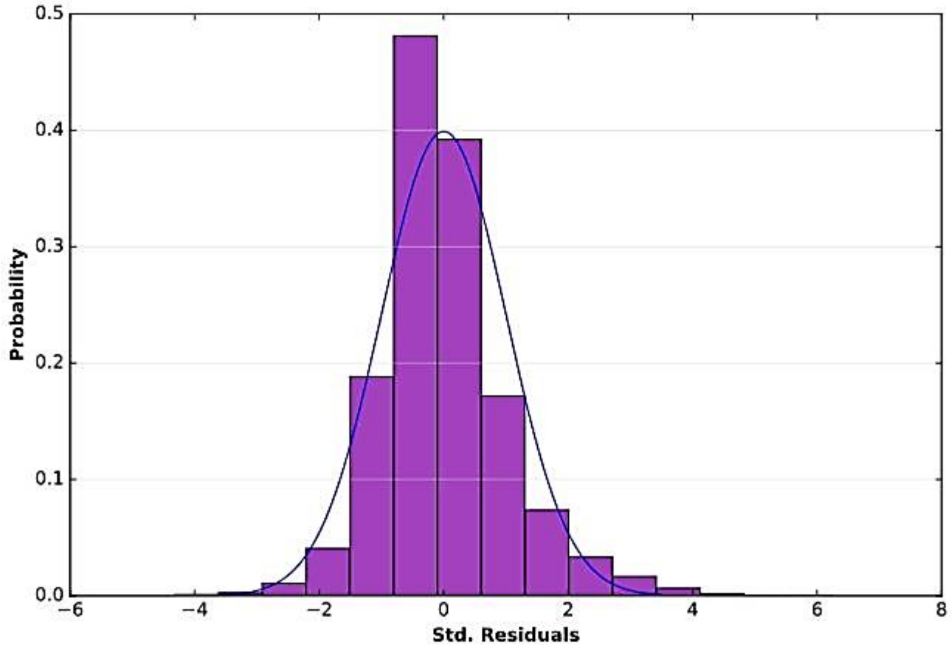



Figure 5.5: OLS4 histogram of standardized residuals.

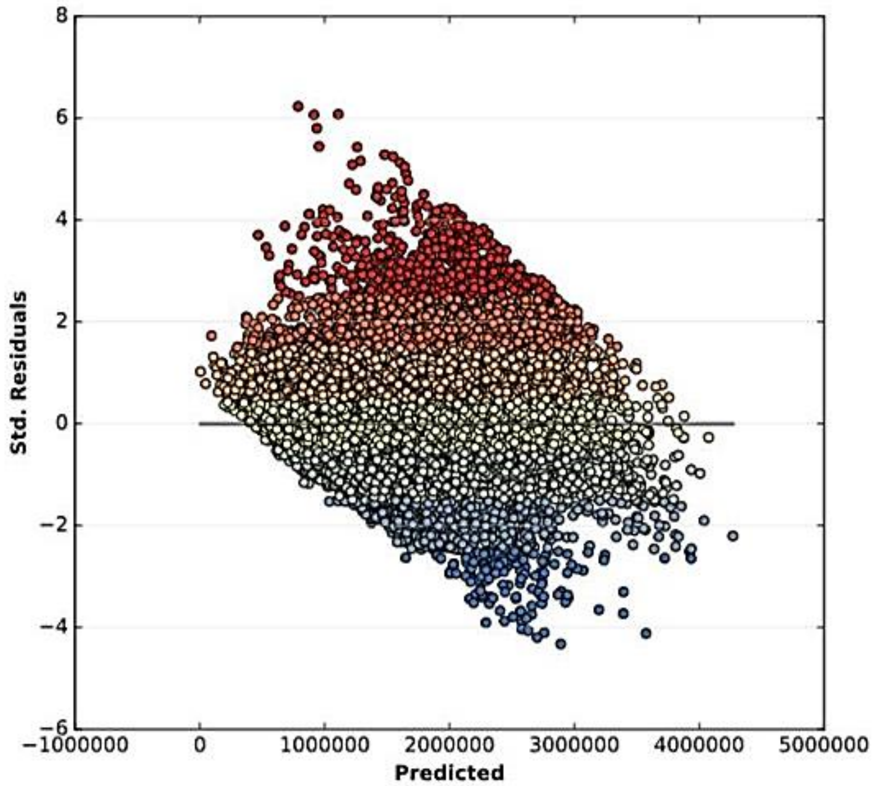



Figure 5.6: OLS4 residuals vs predicted plot.

### 5.4.3 Final GWR (GWR2) results

The results indicated that the GWR2 model using 100 neigbors as bandwidth the adj $R^2$ were found to be 0.80. AIC value resulting from the GWR2 model is much more smaller compared to the one resulted from applying OLS4 which indicates the importance for moving from a global to a local regression modeling. Figures 5.7 and 5.8 illustrates the spatial pattern of residuals in both the conventional OLS4 and the GWR2 models, revealing that GWR2 residuals are notably closer to zero in comparison to OLS4 residuals. Applying less than 100 neigbors as a bandwidth were too small to compute results, indicating that the geographically weighted regression (GWR) algorithm could not find a sufficient number of neighboring data points within the specified distance or with the given number of neighbors to perform the analysis. This occured due to the overly small distance parameter since some of the housing transaction have the same location as they have been moved to the center of the closest building. Applying bandwidths around 100 gives $R^2$ values that are very close to 0.88 while applying bandwidths that are higher than 150 produces $R^2$ values that are close to 1, which means that we can consider 100 bandwidth as the optimum limit to which the model can give the most logical results when examining the connection between the explanatory variables and the distribution of the independent variable.

Removing outliers led to a model with a better fit (as indicated by the lower AIC of GWR2 equaling 278.64 compared to AIC from GWR1 which equals 301.85 but lower explanatory power as indicated in Table 5.6 by the lower adj $R^2$ of GWR2 that equals 0.80 compared to adj $R^2$ from GWR1 which equals 0.86. This suggests that while the model without outliers is more accurate in terms of overall fit, the outliers were contributing significantly to the explained variance in the original model.

In conclusion, the impact of outliers on the Geographically Weighted Regression (GWR) model for housing prices was significant. The initial model, which included outliers, exhibited a high adj $R^2$ of 0.86, suggesting a strong explanatory power. However, after removing outliers, the adj $R^2$ decreased to 0.80, indicating a reduction in the explained variance. Despite this decrease, the Akaike Information Criterion (AIC) improved from 301 to 278, reflecting a better model fit.

These results underscore the importance of addressing outliers in regression analyses. While the model with outliers appeared to have greater predictive power, it was potentially misleading due to the undue influence of extreme values. The model without outliers, though showing lower explanatory power, is more robust and likely to be more generalizable. Future research should consider the role of outliers and apply robust methods to ensure accurate and reliable results.

Figure 5.7 and Figure 5.8 provide a comparative analysis of the residuals distribution for the OLS4 and GWR2 models. Figure 5.8 shows the map of the standard deviation of residuals distribution for the GWR2 model. It is evident from Figure 5.8 that the GWR2 model's residuals are more closely clustered around zero ( Ivory colored dots) compared to the OLS4 model Figure 5.7. This indicates that the GWR2 model provides a better fit, with fewer large residuals and a more accurate representation of the spatial variability in the data.

Table 5.6: GWR2 results using 100 neighbors.

| Var_Type | VAR | R2 | R2Adj | AICc | Neighbors |
|---|---|---|---|---|---|
| **Dependent Field** | Price | 0.88 | 0.80 | 278.64 | 100 |
| **Explanatory Field** | Area /m2 | | | | |
| **Explanatory Field** | No_of_room | | | | |
| **Explanatory Field** | Floor | | | | |
| **Explanatory Field** | NEAR_DIST_UNIVERSITY | | | | |
| **Explanatory Field** | NEAR_DIST_HOSPITAL | | | | |
| **Explanatory Field** | NEAR_DIST_SCHOOL | | | | |
| **Explanatory Field** | NEAR_DIST_TRANSPORTATION | | | | |
| **Explanatory Field** | NEAR_DIST_BANK | | | | |
| **Explanatory Field** | NEAR_DIST_MIGROS SUPERMARKET | | | | |
| **Explanatory Field** | NEAR_DIST_BUDGET SUPERMARKET | | | | |

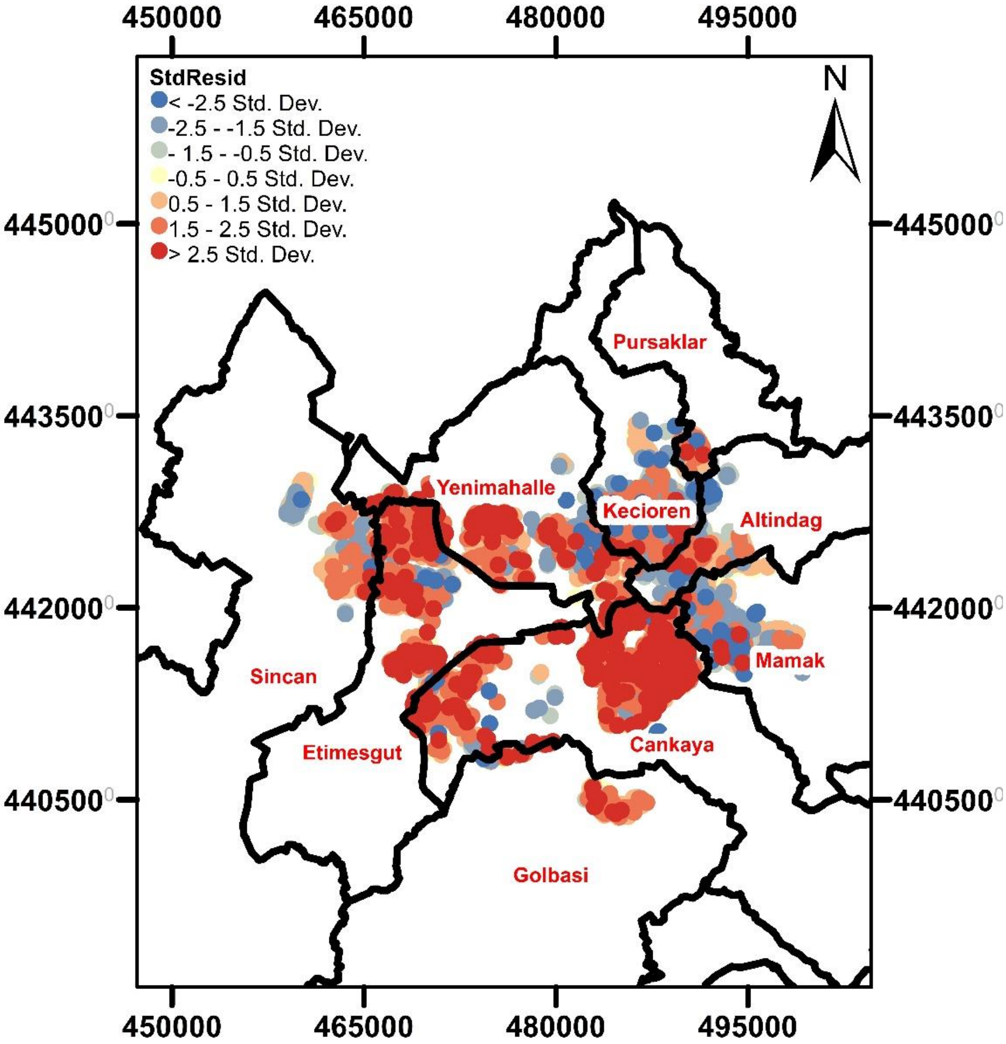



Figure 5.7: OLS4 model residuals.

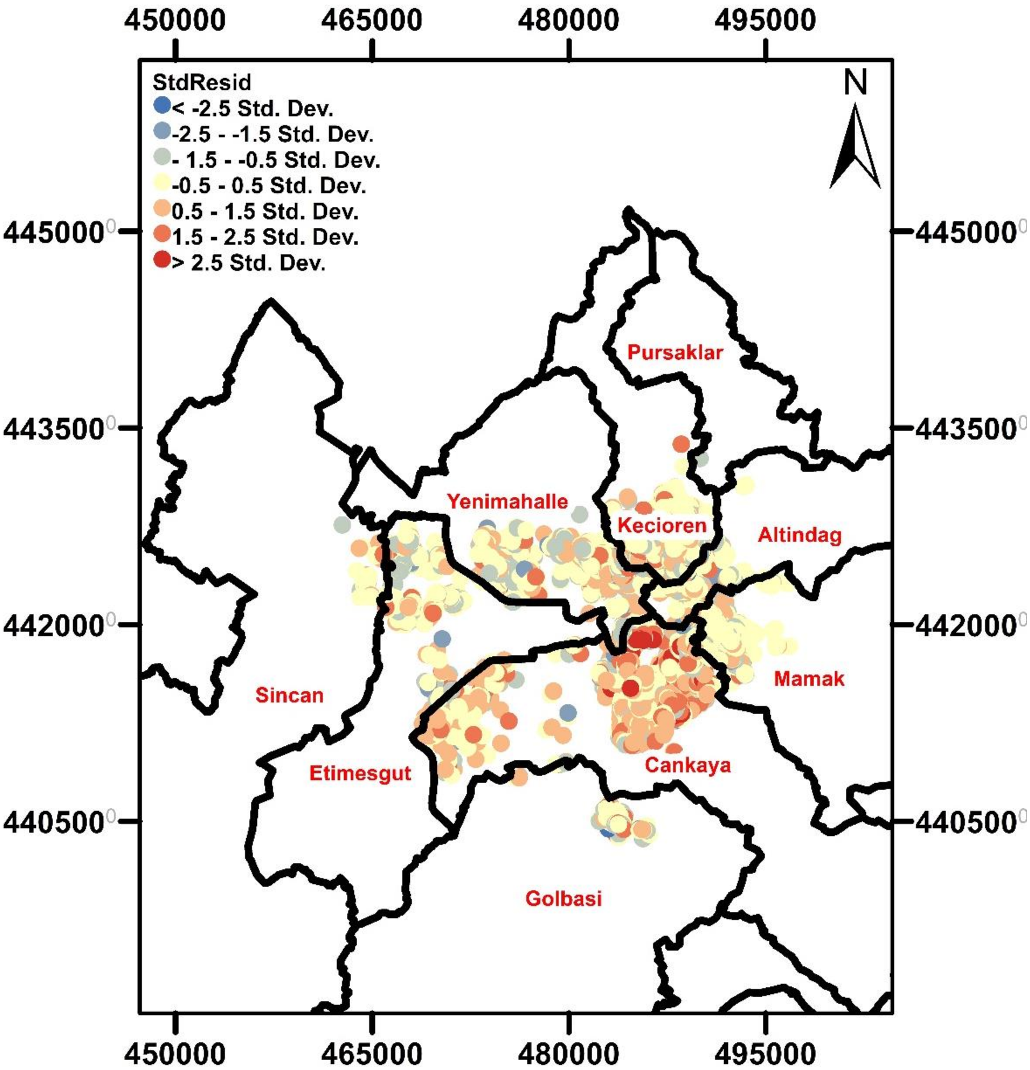



Figure 5.8: GWR2 model residuals.

## 5.5 District Level GWR model (GWR3&GWR4) results

In this study, the dataset was separated into two groups, Çankaya and the rest of the districts, due to Çankaya's unique characteristics. As the central district of Ankara with higher housing prices driven by its commercial and business facilities, Çankaya required a distinct analysis. This allowed for a more accurate understanding of spatial heterogeneity in the housing market.

The initial OLS model applied to the entire dataset showed non-stationarity relationships and spatial heterogeneity, as indicated by the Koenker (BP) test. Therefore, GWR was applied instead of OLS for the separated datasets, offering better insights into local variations.

For the full dataset, GWR2 produced an adjusted $R^2$ of 0.80 and an AIC of 278.64, indicating a strong model fit. When applied separately, the GWR3 model for Çankaya had an adjusted $R^2$ of 0.71 and an AIC of 1101.3, reflecting the complexity of the local market. For the rest of the districts, the GWR4 model achieved an adjusted $R^2$ of 0.60 and an AIC of 10748, suggesting a more homogeneous spatial influence on housing prices.

Overall, GWR proved superior in capturing local variations, highlighting the importance of spatial heterogeneity in housing price analysis. The differences in adjusted $R^2$ and AIC values between the whole and separated datasets reflect varying regional dynamics. Maps of the standard deviation of residuals for these models are presented in Appendix I.

## 5.6 Housing Price Analysis GIS Application Results

The GIS application developed in this study includes several tools that facilitate the analysis and visualization of housing prices and their relationship with various explanatory factors across Ankara's metropolitan area. These tools enable users to understand the heterogeneity in housing prices from one region to another and allow for further analysis of the factors influencing these price variations.

### 5.6.1 Comparative analysis of Kızılay and Çayyolu

Using the GIS application, a comparative analysis of two neighborhoods, Kızılay and Çayyolu, was conducted to assess how various amenities influence housing prices across five price ranges in Ankara. The results highlight significant differences between the two neighborhoods, particularly in terms of the availability of schools and transportation, which appear to contribute to the price variation between Kızılay and Çayyolu. Table 5.7 presents the housing price ranges classified in five classes. The number of ameneties within easch of these ranges for both Kızılay and Çayyolu is shown in Tables 5.8.

Table 5.7: Housing price ranges in Ankara.

| Price Range | Price Interval (USD) |
|---|---|
| PRange1 | 17,280 – 45,360 |
| PRange2 | 45,360 – 63,180 |
| PRange3 | 63,180 – 88,560 |
| PRange4 | 88,560 – 129,060 |
| PRange5 | 129,060 – 218,700 |

Table 5.8: Distribution of amenities in Kızılay across five price ranges and Çayyolu across two price ranges.

| Price Range | Kızılay | | | | | Çayyolu | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | PRrange1 | PRrange2 | PRrange3 | PRrange4 | PRrange5 | PRrange1 | PRrange2 | PRrange3 | PRrange4 | PRrange5 |
| **Number of Malls** | 1 | 0 | 0 | 0 | 0 | --- | --- | --- | 0 | 0 |
| **Number of Schools** | 0 | 0 | 2 | 3 | 0 | --- | --- | --- | 2 | 22 |
| **Number of Universities** | 0 | 0 | 0 | 0 | 0 | --- | --- | --- | 0 | 0 |
| **Number of Hospitals** | 0 | 0 | 0 | 0 | 0 | --- | --- | --- | 0 | 0 |
| **Number of ATMs** | 41 | 9 | 47 | 42 | 14 | --- | --- | --- | 7 | 8 |
| **Number of Banks** | 9 | 0 | 16 | 14 | 4 | --- | --- | --- | 0 | 2 |
| **Number of Budget Supermarkets** | 0 | 0 | 0 | 2 | 1 | --- | --- | --- | 0 | 4 |
| **Number of Supermarkets** | 1 | 0 | 0 | 0 | 0 | --- | --- | --- | 1 | 3 |
| **Number of Transportation Points** | 6 | 0 | 18 | 16 | 0 | --- | --- | --- | 25 | 55 |

### 5.6.2 Interpretation of results

The comparative analysis of Kızılay and Çayyolu reveals important distinctions in terms of amenity distribution, which correspond to housing price variations in these neighborhoods. In Kızılay, housing transactions are spread across all five price ranges. However, the number of key amenities such as malls, schools, and transportation points varies significantly. For instance, in the third and fourth price ranges, the presence of schools and transportation is more prominent, likely contributing to higher prices in these ranges. The availability of ATMs and banks is also concentrated in these middle ranges, further supporting the idea that the clustering of amenities plays a role in driving prices up.

In Çayyolu, transactions are only found in the fourth and fifth price ranges, indicating that housing prices are generally higher compared to Kızılay. This is reflected in the abundance of schools and transportation points in the fifth price range, with 22 schools and 55 transportation points, which stand in stark contrast to the lower-priced areas of Kızılay. The presence of more budget supermarkets and supermarkets in Çayyolu's higher price ranges also suggests that these amenities serve to a different segment of the housing market, further pushing prices up in this region.

Figures 5.9-5.12 illustrate the differences in the number of schools and transportation points between Kızılay and Çayyolu, highlighting how these factors are concentrated in areas with higher housing prices. The strong correlation between transportation access and schools with higher price ranges, particularly in Çayyolu, suggests that these amenities play an important role in determining housing prices in Western Çankaya. Additionally, these findings may be linked to the specific characteristics of each neighborhood. Kızılay, for example, is a local business center, which could explain the lower presence of schools and certain other amenities compared to more residential areas like Çayyolu. The business-centric nature of Kızılay may contribute to a different distribution of amenities, focusing more on commercial services such as banks, and ATMs rather than schools and residential infrastructure. This suggests that the amenity-price relationship seen here may be unique to these neighborhoods, and results could vary in other areas with different characteristics. Therefore, in other regions, factors beyond those considered in this study may need to be analyzed and incorporated into the GIS application, taking into account the specific nature and purpose of the area being studied. This flexibility in modeling will ensure that the application can adapt to the unique drivers of housing prices in diverse urban environments.

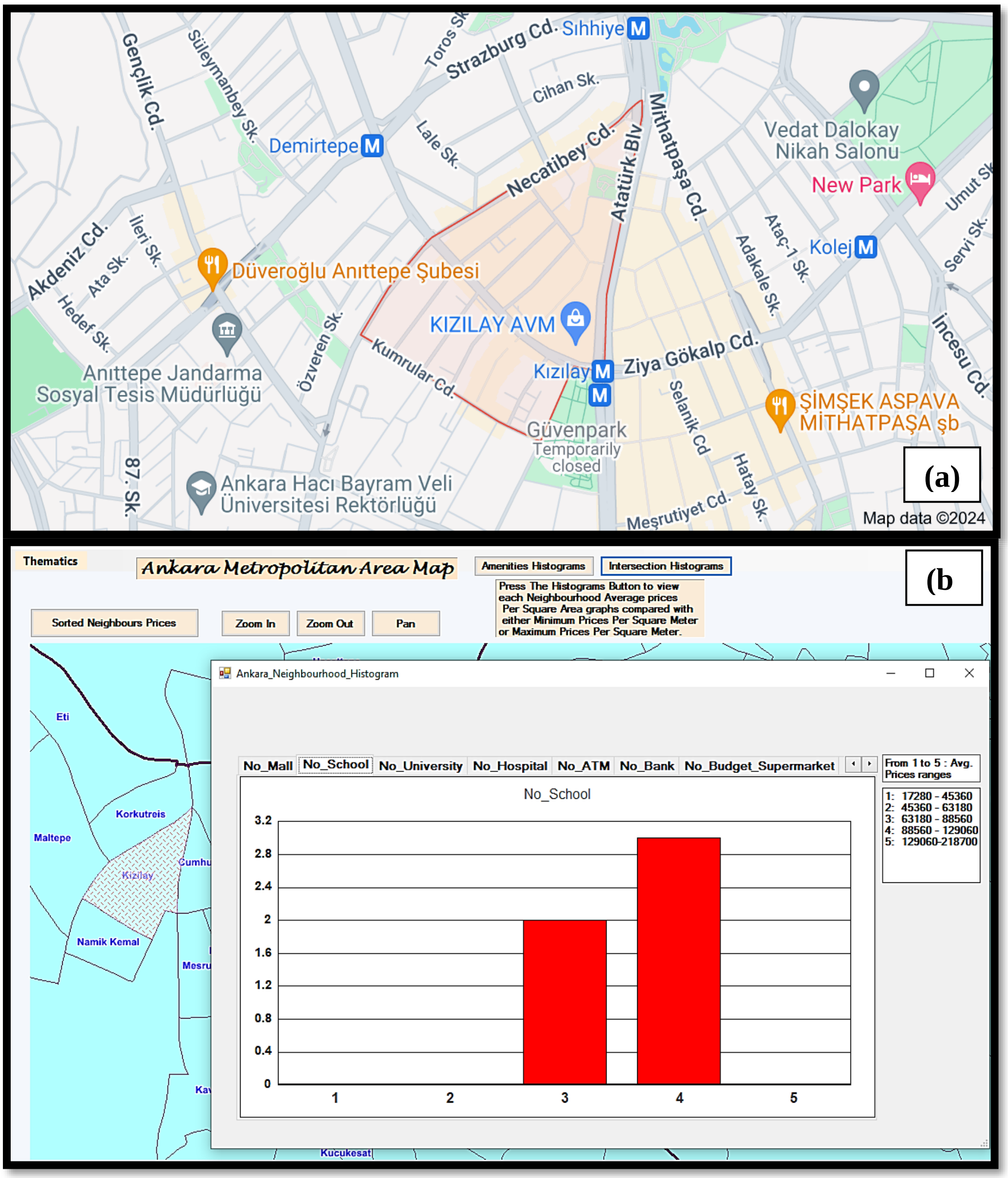



Figure 5.9: a) Kızılay neighborhood (URL15). b) Distribution of Schools in Kızılay within five price ranges.

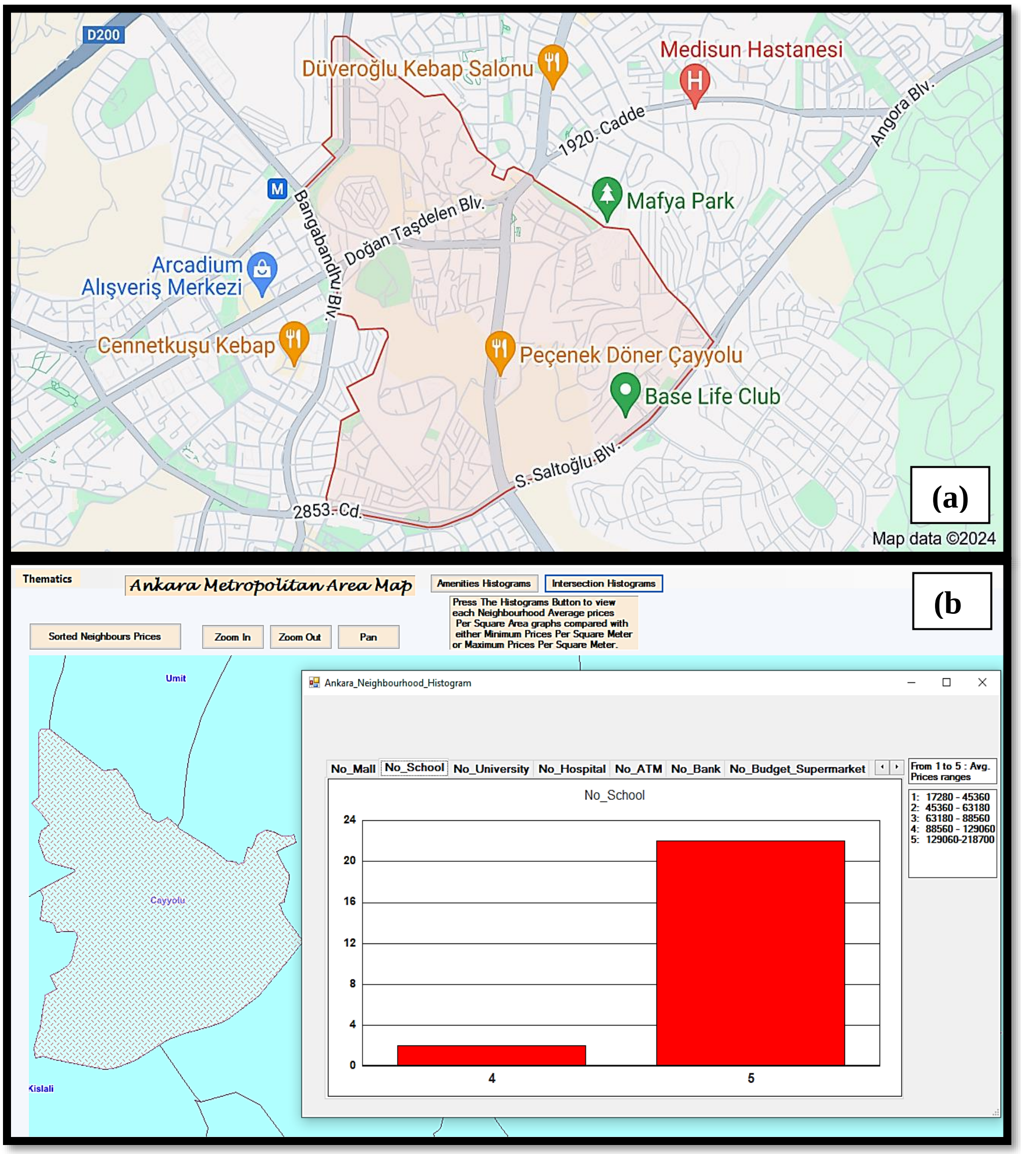



Figure 5.10: a) Çayyolu neighborhood (URL16). b) Distribution of Schools in Çayyolu within five price ranges.

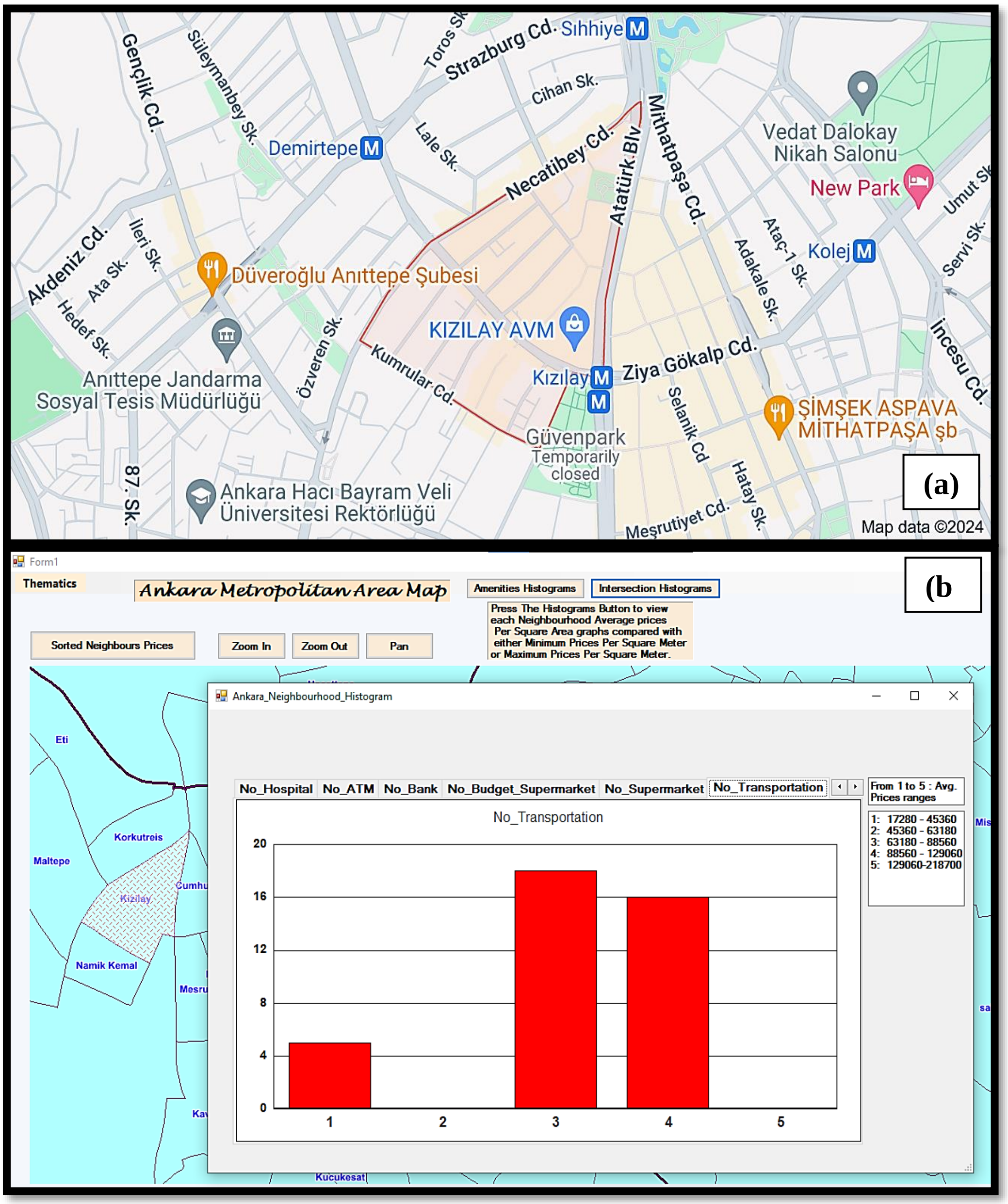



Figure 5.11: a) Kızılay neighborhood (URL15). b) Distribution of Transportation Points in Kızılay within five price ranges.

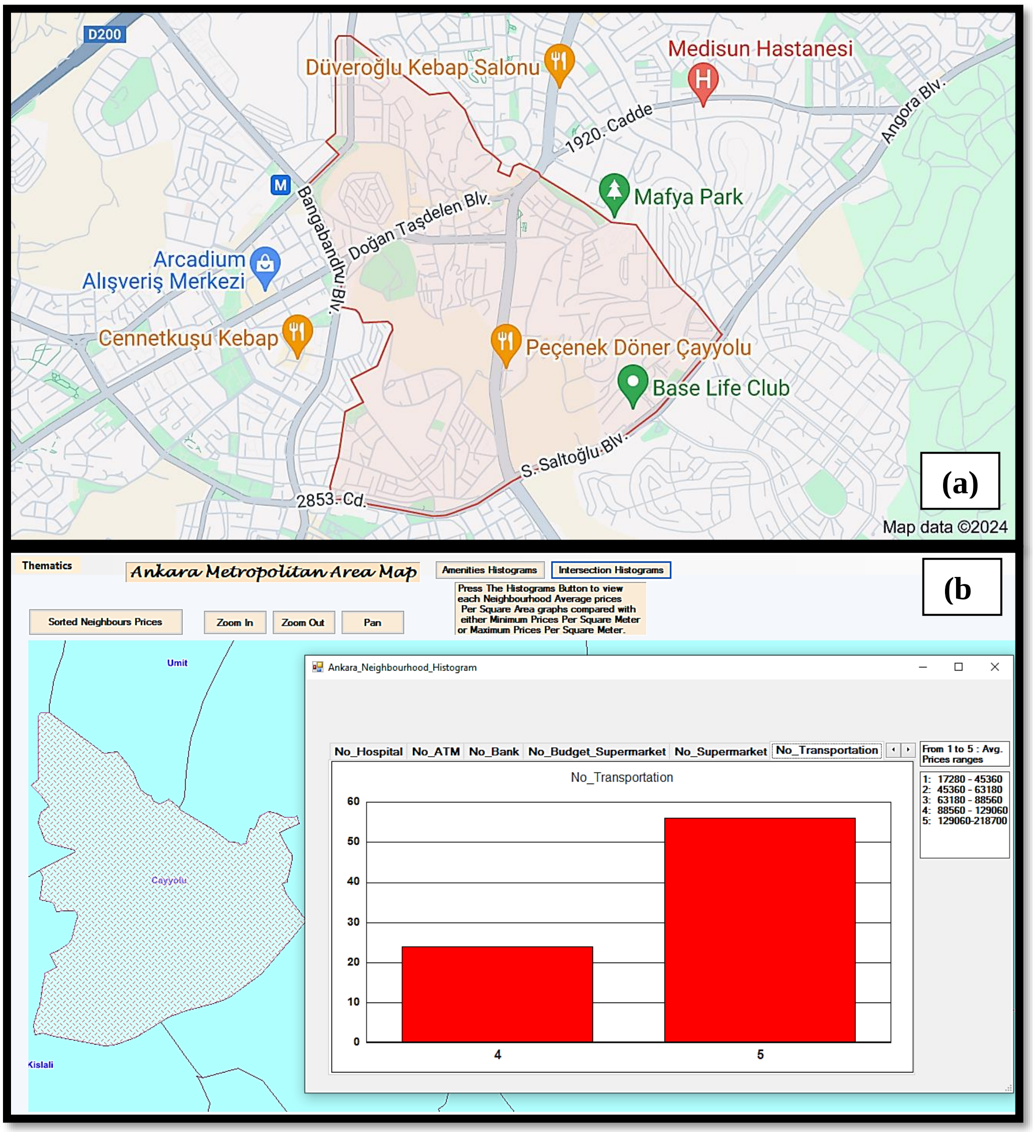



Figure 5.12: a) Çayyolu neighborhood (URL16). b) Distribution of Transportation Points in Çayyolu within five price ranges.

Figure 5.13 demonstrates how the GIS application effectively maps results and illustrates their relationships in real-time. In this instance, the thematic map of Kızılay showcases five housing price ranges, overlaid with symbols representing transportation and schools. When observing the fifth price range, which represents the highest housing prices in the area, it is notable that there are no transportation or school points directly within this zone.

This application highlights that, on a broader district comparison level, higher price ranges typically correlate with more amenities. However, at a more localized level, the situation can differ. In this case, while transportation and school points are not directly within the highest price range, they are situated very close to it. These results are influenced by the way the Thiessen zones were constructed, which determines the spatial distribution of these amenities relative to housing prices.

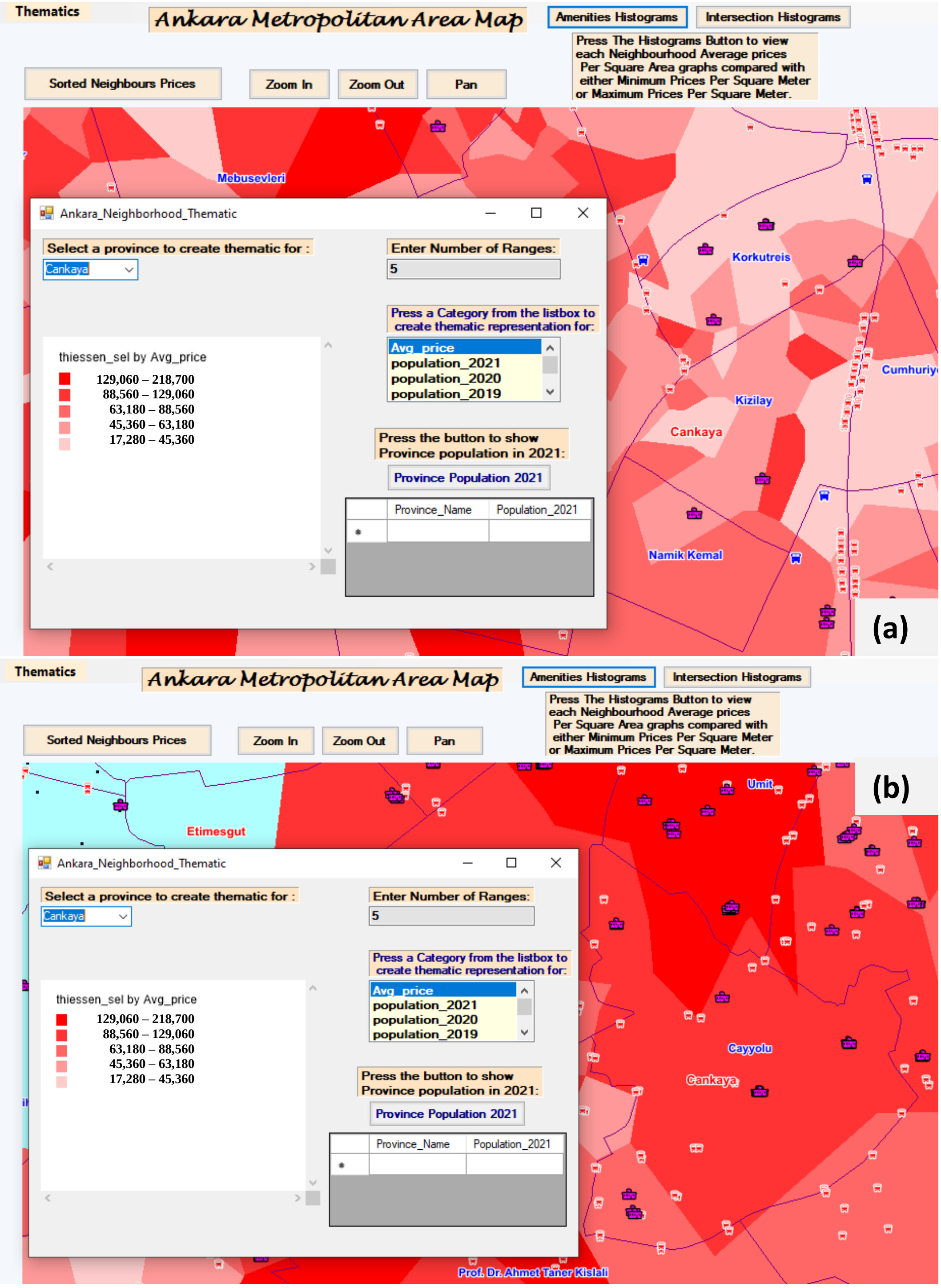



Figure 5.13: Thematic Maps of Kızılay (a) and Çayyolu (b) Neighborhoods with Thiessen Polygons, Schools & University amenities, and Price Ranges.

# CHAPTER 6

# DISCUSSION OF THE RESULTS

This study explores the examination of a pivotal question regarding the potential existence of spatial heterogeneity in housing prices within the Ankara metropolitan area, exploring the impact of various factors on these price distributions. The findings of this research affirm that specific factors, including the proximity of homes to universities, malls, hospitals, and transportation, along with essential house characteristics like the number of rooms, and floor level, exert a significant influence on determining housing prices. The discernible spatial heterogeneity in this price distribution underscores the intricacy of the real estate market dynamics. The following subsections discusses the workflow of this study.

## 6.1 Data Preperation and Processing Challenges

The process of obtaining the data for this study involved significant challenges, particularly in retrieving information from the Sahibinden real-estate website. The data was initially extracted as a JSON file, which required extensive conversion into more usable formats such as Excel and shapefiles. This transformation process was time-consuming and complex, underscoring the difficulties often encountered in real estate data acquisition. The struggles in converting the data were compounded by the need to ensure accuracy and completeness, vital for the subsequent spatial analysis. As highlighted by Goodchild and Li (2012), advanced spatial analysis tools are important in capturing the complexity of geographic phenomena, making the accuracy of initial data retrieval and preprocessing a critical step.

Preprocessing the dataset was an important step to ensure the reliability of the analysis. This phase involved cleaning the data by removing rows with missing information and identifying and eliminating outliers. The exclusion of these outliers was particularly important, as they could distort the results of the regression models. Through careful data preprocessing, the integrity of the dataset was maintained, enabling a more accurate analysis of the spatial distribution of housing prices in Ankara. The necessity of this data cleaning process is supported by Dubin (1992), who noted the potential distortions caused by unaddressed data issues in housing market models.

Additionally, the housing prices data underwent a spatial transformation, shifting from its original geographical location based on the Sahibendan site to the center of the nearest adjacent house. This adjustment was necessitated by the prevalence of multi-storey buildings and the availability of price information for various apartments within the same structure, resulting in overlapping data points at the center of the building. This spatial distortion, marked by multicollinearity issues, posed challenges during the application of automatic geographically weighted regression methods. Therefore, when the cross-validation (CV) technique and the Akaike Information Criterion (AIC) were applied for the model to determine the optimal bandwidth depending on the characteristics of the dataset, the presence of collinearity resulted in locality issues in the dataset. To address this challenge, geographically weighted regression (GWR) model were computed using a Gaussian kernel weighting function. The model utilized both an adaptive bandwidth method and a manual bandwidth parameterization method, contingent on a predefined number of neighbors threshold. Had the original data been accurately geolocated, sourced from entities like data systems companies or governmental departments such as the Department of Lands, Survey, and Statistics, the results of the geographically weighted regression could potentially have yielded greater accuracy.

## 6.2 Kriging Geostatistical Technique

The application of the Kriging geostatistical technique provided insights into the spatial correlation of housing prices. The model indicated that spatial dependence starts to level out at a distance of approximately 17 kilometers. This finding is significant, as it suggests that most of the pricing data within the Ankara metropolitan area is considered correlated, given that the diameter of the densest pricing data is around 22 kilometers. This step was essential in understanding the spatial structure of the data before applying regression models.

## 6.3 Application of OLS Models

Four Ordinary Least Squares (OLS) models were applied in this study (OLS1, OLS2, OLS3, OLS4) on the entire dataset to explore the impact of various explanatory variables. The analysis revealed that some variables, such as Area/m² and distance to ATMs, were insignificant in explaining housing prices and were subsequently excluded in OLS2 and OLS4. The removal of outliers between OLS3 and OLS4 highlighted the importance of data refinement in enhancing model accuracy, leading to more robust conclusions about the factors influencing housing prices.

## 6.4 Application of GWR Models

Two Geographically Weighted Regression (GWR) models (GWR1 and GWR2) were applied to account for spatial heterogeneity in the dataset. GWR1 was used before excluding outliers, and GWR2 was applied afterward, with the non-significant variables from the OLS models excluded. The GWR models provided a more localized understanding of how different factors influenced housing prices, offering a spatially detailed perspective that complemented the OLS findings. The emphasis on spatial heterogeneity and local contextual factors in this analysis is supported by studies in urban economics (Leishman, 2009; Machin & Gibbons, 2008).

## 6.5 Comparison of Final Results (OLS4 vs. GWR2)

The final comparison between the results of OLS4 and GWR2 highlighted the strengths and limitations of each approach. While OLS4 offered a broader overview of the factors affecting housing prices, GWR2 provided detailed insights into the spatial variability of these influences. The comparison underscored the importance of considering spatial heterogeneity in real estate analyses, as it revealed significant variations that a global model like OLS might overlook.

The analysis suggests that the GWR2 (Geographically Weighted Regression after outlier removal) model is the most effective in explaining variations in housing prices within the Ankara metropolitan area. This model outperforms the OLS models by capturing spatial heterogeneity, reflecting the varying influence of explanatory variables across different locations.

The application of Geographically Weighted Regression (GWR) in this study highlighted the spatial heterogeneity of factors influencing housing prices across different areas of Ankara. This aligns with findings from Apaydin & Güneş, (2022), who also observed that the impact of various explanatory variables on housing prices varies significantly across different locations within the city. Specifically, the study found that the effects of variables such as dwelling size, number of bedrooms, and net surface area were particularly strong in specific micro-areas within the Çankaya district. This observation is consistent with the results of this study, where similar variables demonstrated a significant positive impact on housing prices in the same region, including number of rooms and floor number.

Moreover, Apaydin & Güneş, (2022) showed that distance-related variables, including proximity to hospitals and schools, generally had a negative impact on housing prices, especially in areas farther from the central business district. This finding supports the conclusion drawn in this study that proximity to essential services like universities and hospitals is an important determinant of housing prices in Ankara. However, the extent of this impact varies depending on the specific location within the metropolitan area.

These results underscore the limitations of traditional OLS models, which often fail to account for the spatial heterogeneity present in housing markets. In contrast, the GWR model provides a more clear understanding by capturing the local variations in how different factors influence housing prices, making it a valuable tool for urban planners and policymakers.

## 6.6 Choosing the Size of the Model Area

Çankaya, as a distinct district within Ankara, warranted separate analysis due to its unique characteristics as a commercial and business hub. The study found that the factors influencing housing prices in Çankaya differed significantly from those in the rest of the metropolitan area. By applying GWR models exclusively to Çankaya and the rest of the districts separately, the analysis highlighted the district-specific dynamics that are important for localized urban planning and real estate decision-making. In contrast to the research conducted by Apaydin & Güneş, (2022), this study underscores the significance of Gölbaşı district, prompting its inclusion in the scope of investigation within the Ankara metropolitan area. The decision to include Gölbaşı district, was driven by its significant transaction volume, reinforcing the importance of considering diverse districts in metropolitan housing market studies. As well as motivated by the notable share of transactions occurring in Gölbaşı district, surpassing the threshold of 3%. Previous studies have highlighted the importance of including diverse districts in metropolitan analyses to understand regional variations and market dynamics better (Bitter et al., 2007).

## 6.7 Development of the Customized GIS Application

The development of a customized GIS application represents a significant contribution of this study, serving as a powerful tool for visualizing and analyzing housing price distributions across Ankara's metropolitan area. This application leverages advanced spatial analysis techniques, including heat maps and Thiessen polygons, to offer a better understanding of how housing prices vary across different districts and neighborhoods. By integrating these features, the GIS application provides a detailed and user-friendly platform for examining the spatial heterogeneity of the housing market, making it an invaluable resource for various stakeholders.

- A key strength of the customized GIS application is its ability to provide advanced spatial analysis through heat maps, which visually represent housing price concentrations across the metropolitan area. These maps reveal patterns that may not be visible through traditional statistical methods, highlighting high and low price areas. This allows users to quickly identify key regions, such as high-priced hotspots or lower-priced zones, making the tool invaluable for urban planning and policy-making.
- The integration of Thiessen polygons enhances the GIS application by dividing space into distinct regions based on proximity to housing transaction points. This method helps clearly define areas of influence for each transaction, offering a more accurate representation of how neighborhoods are impacted by local housing dynamics. By providing precise spatial delineation, it allows for a detailed understanding of localized housing price variations and the factors driving these differences.
- The GIS application developed in this study has practical uses for urban planners, real estate developers, and researchers. For urban planners, it enables spatial visualization of housing price distributions, aiding in decisions about zoning, infrastructure, and resource allocation. Planners can identify areas needing investment for balanced development. Real estate developers can use the tool to pinpoint lucrative investment areas, assessing how proximity to amenities like schools and hospitals impacts housing prices, which helps guide strategic decisions on property development.
- For researchers, the GIS application offers a powerful tool to study factors influencing housing prices by analyzing variables like amenity proximity and population density. Its visual and interactive features help explore complex relationships, generate insights, and refine theories on urban economics and housing markets, enabling deeper, more comprehensive analysis.

- The GIS application's user-friendly interface makes it accessible to users with limited GIS experience. Its intuitive design, interactive tools, and simplified spatial analysis encourage wider use among stakeholders, promoting data-driven decision-making in urban planning and real estate development.
- The GIS application provides valuable insights for policy-making, particularly in addressing housing inequality and affordability. By highlighting high-priced areas and their contributing factors, it helps policymakers craft strategies for increasing affordable housing and improving service distribution, promoting more equitable urban development.

## 6.8 Future Improvement on Analysis Subzoning

The current GIS application lays the groundwork for future improvements, such as adding real-time data tracking and machine learning to predict housing trends. It can also be scaled to other cities, extending the insights and methodologies from this study to broader urban contexts, particularly when collaborating with experts in urban planning. The insights gained from this analysis could then be more effectively implemented in urban policy and development strategies. Enhancements might include road networks zoning concepts that deepen global understanding of housing market dynamics.

In addition to analyzing the effect of various amenities on housing prices at both the district and neighborhood levels, this study also considers the potential influence of main road-closed zone areas on property values. These zones may consist of one or multiple neighborhoods, might also be spread across several districts, that are defined by the boundaries of major roads. where Zone A lies in Çankaya, Zone B in Etimesgut, and Zone C in Mamak. The connectivity of these roads can either include these zones or form the borders of zones that may share common characteristics, which could influence housing prices in a way distinct from picking random individual neighborhood or on district-level factors.

The hypothesis is that different types of amenities, such as transportation hubs, schools, commercial centers, and green spaces, may need to be analyzed based on the unique properties of these picked zones based on main roads. The nature of these zones—whether they are well-connected, isolated, or predominantly residential or commercial—could play an important role in shaping housing price trends. For instance, Zone A, within Çankaya, may exhibit higher property values due to its proximity to main roads and central amenities, whereas Zone B in Etimesgut might experience different valuation effects due to its suburban context and distinct road access patterns. Additionally, other factors related to road infrastructure, such as traffic patterns could further shape housing values in these zones.

To illustrate this concept, Figure 6.1 sets the foundation for understanding how major road networks serve as natural boundaries for residential zones in central Ankara. By delineating each zone, it becomes easier to analyze the influence of road connectivity on housing characteristics within these zones. This overview is essential to set the scene for the subsequent figures, which detail each zone individually. However, due to the lack of access to detailed road network data in the Ankara metropolitan area, the full potential of this type of spatial analysis remains unexplored. Future research involving urban planning specialists can understand deeper the relationship between main road networks and housing prices once the necessary data becomes available, offering a more comprehensive understanding of how road connectivity and zoning patterns affect property values across the city.

Figure 6.1 provides an overview of three selected residential zones in Ankara: Zone A (Bahçelievler-Yukarı Bahçelievler), Zone B (Batı Sitesi-Ergazi), and Zone C (Aşık Veysel-Abidinpaşa). Each zone is labeled and depicted on a base road map of Ankara, highlighting their location and proximity to major main roads. This map gives a contextual view of how these residential zones are defined by surrounding road networks, making the concept of main road-defined zones visible.

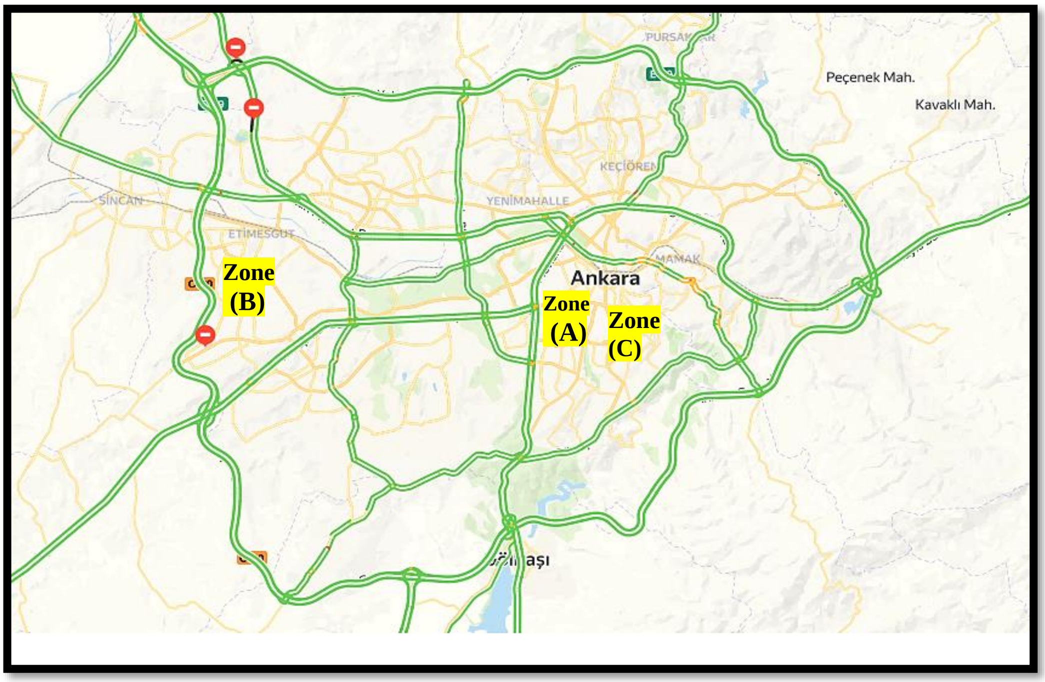



Figure 6.1: Ankara metropolitan area main roads map (URL17).

Figure 6.2 presents the full view of Zone A, comprising Bahçelievler and Yukarı Bahçelievler neighborhoods. The area is shown as a designated residential land-use zone, bounded by main roads, with labels indicating the neighborhood names. Zone A is bordered by İsmet İnönü Blulvarı on the south, Mareşal Fevzi Çakmak Caddesi on the east, Mevlana Blulvarı on the west, and Bahriye Üçok Caddesi on the north. This depiction highlights Zone A's residential structure within a framework of key arterial roads, demonstrating the impact of these roads on defining neighborhood characteristics and accessibility.

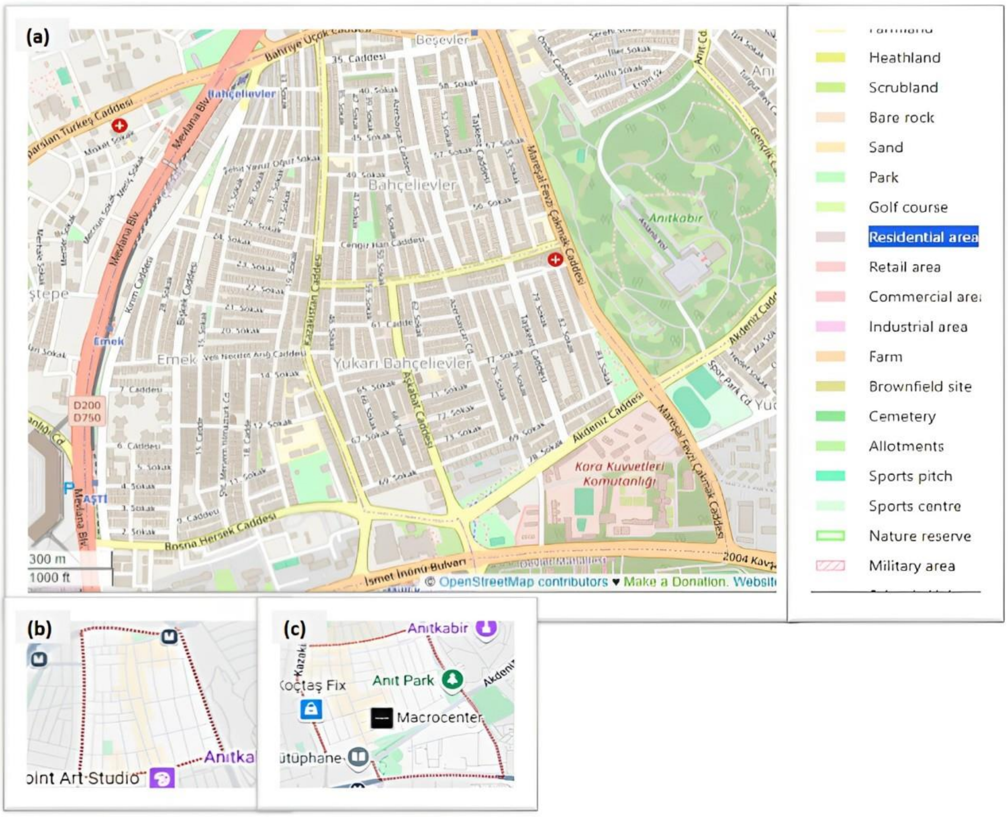


Figure 6.2: a) Zone A - Entire Area. b) Bahçelievler Neighborhood Boundaries. c) Yukarı Bahçelievler Neighborhood Boundaries (URL18).

Zone B is shown as a residential area in Figure 6.3 encompassing the Batı Sitesi and Ergazi neighborhoods. The zone's boundaries are outlined by major roads, with neighborhood names labeled within. The zone is framed by Fatih Sultan Mehmet Bulvarı to the south, Batıkent Bulvarı to the east, and Cengiz Aytmatov Caddesi to the north. Zone B's layout illustrates the influence of main roads on suburban residential areas, where proximity to both industrial and residential zones affects accessibility and housing patterns.

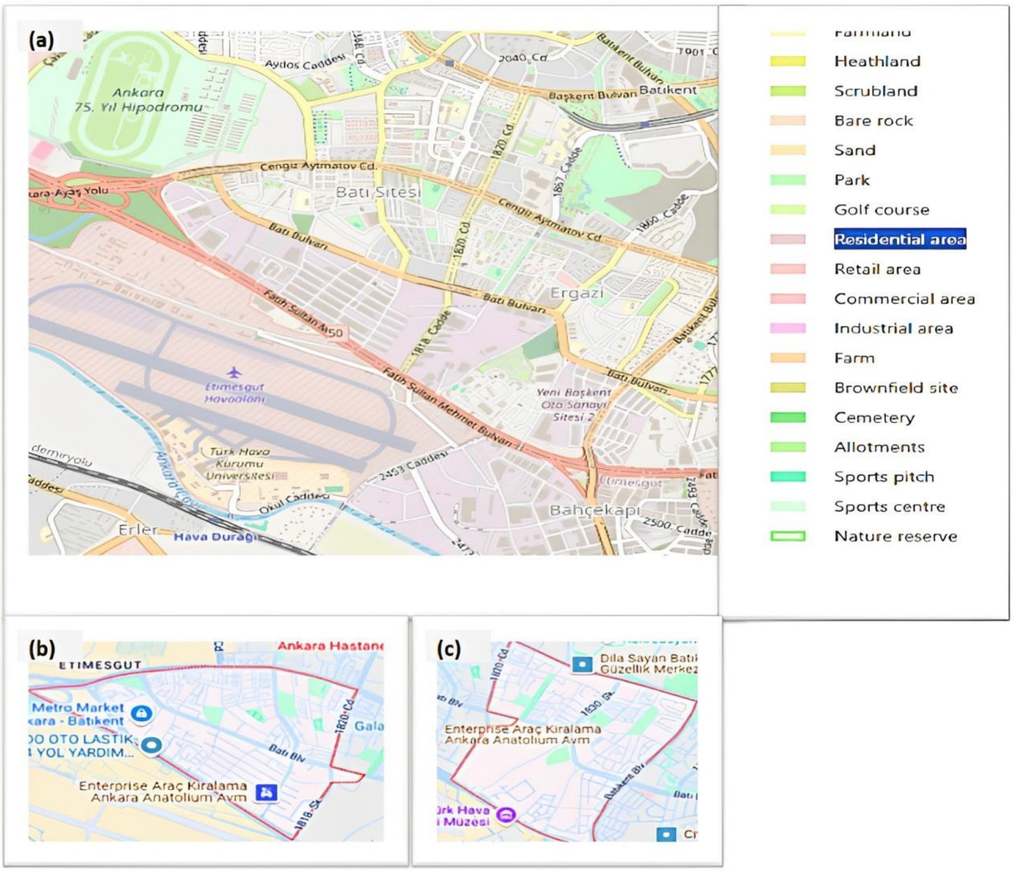



Figure 6.3: a) Zone B - Entire Area. b) Batı Sitesi Neighborhood Boundaries. c) Ergazi Neighborhood Boundaries (URL18).

Figure 6.4 displays the overall Zone C, encompassing Aşık Veysel and Abidinpaşa neighborhoods within a residential area, bordered by major roads and with neighborhood labels. This zone is defined by Tıp Fakültesi Caddesi to the north, Nato Yolu Caddesi to the east, Münzeviler Caddesi to the west, and Mehmet Ali Altın Caddesi to the south. Zone C showcases how outer suburban zones are framed by major roads, allowing for residential development that maintains accessibility to urban centers. This setup offers insights into how main roads influence suburban housing dynamics and pricing.

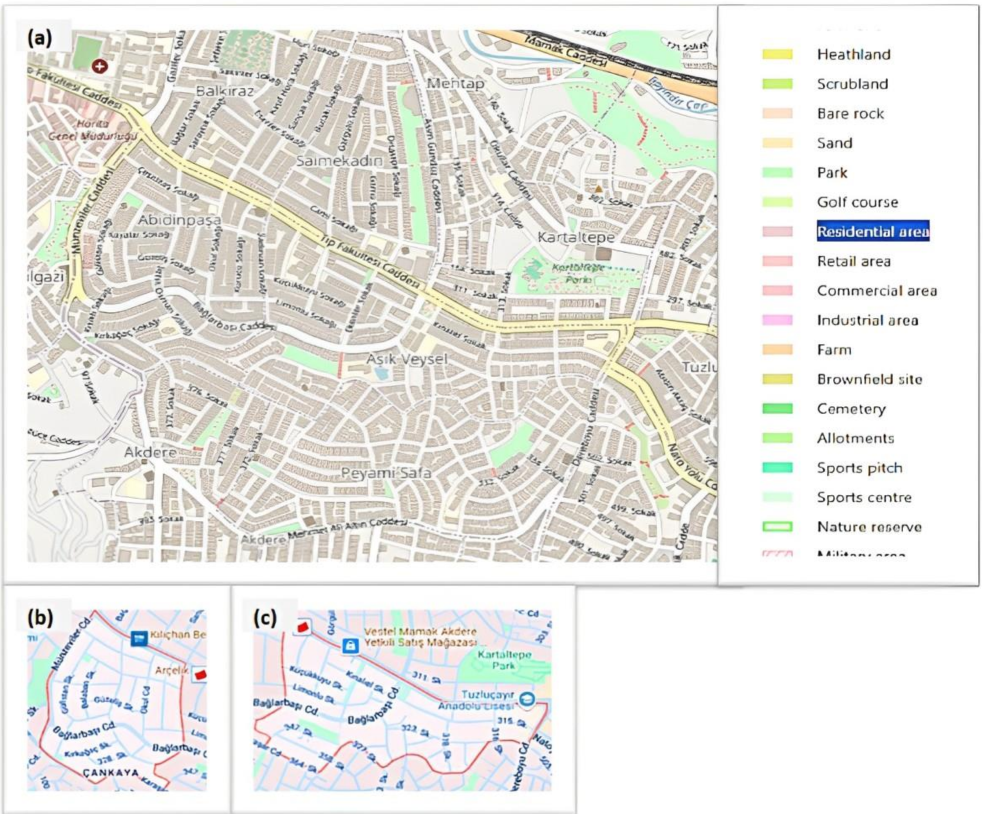



Figure 6.4: a) Zone C - Entire Area. b) Abidinpaşa Neighborhood Boundaries. c) Aşık Veysel Neighborhood Boundaries (URL18).

Such analysis, further emphasize how different configurations of road based zones can offer insights into how amenities, connectivity, and spatial layout impact the housing market. This approach provides a more tailored understanding of property values that go beyond typical district or neighborhood-level analysis.

# CHAPTER 7

# CONCLUSIONS AND RECOMMENDATIONS

This study centers on the development and implementation of a customized GIS application, designed to explore the spatial heterogeneity in housing prices within Ankara's metropolitan area. By integrating Geographically Weighted Regression (GWR) and Ordinary Least Squares (OLS) models, the application provides an interactive platform to analyze how key factors such as proximity to amenities (universities, schools, malls, hospitals, supermarkets, budget supermarkets, banks, ATM and transportation) and house characteristics (number of rooms, floor level, and area) influence housing prices.

The GIS tool developed as part of this research allows users to visualize the spatial distribution of these factors in relation to housing prices, offering a detailed view of how different variables shape real estate dynamics across different districts. This makes the application highly valuable for urban planners, policymakers, and real estate professionals, providing actionable insights to address spatial variations in housing markets.

The analysis confirmed the presence of significant spatial heterogeneity in the data, with GWR and OLS models revealing detailed differences in how factors affect housing prices across locations. Importantly, removing outliers from the dataset enhanced the accuracy of these models, further validating the role of spatial data preprocessing.

In recognition of Çankaya's distinct characteristics as a commercial hub with elevated housing prices, the study also divided the data into two groups—Çankaya and the remaining districts—demonstrating the importance of localized analysis. The GIS application efficiently facilitated this distinction, making it an effective tool for highlighting unique patterns in different regions of Ankara.

Ultimately, the GIS application stands as the primary contribution of this study, providing a powerful method for examining and addressing the complexities of spatial heterogeneity in urban real estate markets. This interactive tool represents a significant advancement in housing price analysis, and its continued use can guide future efforts in infrastructure development, zoning, and policy planning in Ankara and similar metropolitan areas.

This study demonstrates the potential of GIS applications in examining the influence of spatial variables, including the role of main roads in shaping residential zones, on housing prices. The road connectivity may significantly impact property values by affecting neighborhood accessibility and proximity to amenities. Future research with enhanced data on traffic patterns and road networks could deepen these insights, especially if guided by urban planning experts. Such developments would provide a more comprehensive understanding of how road infrastructure and zoning contribute to housing market dynamics.

# REFERENCES

Akkay, R. C. (2021). The Macroeconomic Determinants of The Housing Prices in Turkey. *Erciyes Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, *58*, Article 58. https://doi.org/10.18070/erciyesiibd.801319

Altuzarra, A., & Esteban, M. (2011). Land prices and housing prices: The case of Spain. *Journal of Housing and the Built Environment*, *26*(4), 397–409. https://doi.org/10.1007/s10901-011-9235-8

Apaydin, A., & Güneş, T. (2022). Spatial heterogeneity in housing market: Ankara metropolitan area. *Erciyes Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*. https://doi.org/10.18070/erciyesiibd.1122568

Baba, K., Bahi, L., & Latifa, O. (2015). Geostatistical Analysis for delineating sterile inclusions in Sidi Chennane' phosphatic series, Morrocco. *Earth Sciences Research Journal*, *18*. https://doi.org/10.15446/esrj.v18n2.41173

Bagheri, B., & Shaykh-Baygloo, R. (2021). Spatial analysis of urban smart growth and its effects on housing price: The case of Isfahan, Iran. *Sustainable Cities and Society*, *68*, 102769. https://doi.org/10.1016/j.scs.2021.102769

Baker, E., Bentley, R., Lester, L., & Beer, A. (2016). Housing affordability and residential mobility as drivers of locational inequality. *Applied Geography*, *72*, 65–75. https://doi.org/10.1016/j.apgeog.2016.05.007

Barros, C. P., Chen, Z., & Gil-Alana, L. A. (2013). Long Memory in the Housing Price Indices in China. *Asian Journal of Empirical Research*, *3*(7), Article 7.

Basu, S., & Thibodeau, T. G. (1998). Analysis of Spatial Autocorrelation in House Prices. *The Journal of Real Estate Finance and Economics*, *17*(1), 61–85. https://doi.org/10.1023/A:1007703229507

Bitter, C., Mulligan, G. F., & Dall'erba, S. (2007). Incorporating spatial variation in housing attribute prices: A comparison of geographically weighted regression and the spatial expansion method. *Journal of Geographical Systems*, *9*(1), 7–27. https://doi.org/10.1007/s10109-006-0028-7

Can, A. (1998). GIS and Spatial Analysis of Housing and Mortgage Markets. *Journal of Housing Research*, *9*(1), 61–86.

Cao, K., Diao, M., & Wu, B. (2019). A Big Data–Based Geographically Weighted Regression Model for Public Housing Prices: A Case Study in Singapore. *Annals of the American Association of Geographers*, *109*(1), 173–186. https://doi.org/10.1080/24694452.2018.1470925

Chaudhary, A., Nazir, N., Riaz, A., Sadiq, N., & Riaz, N. (2023). *Autocorrlation: What happens if the error or disturbance terms are correlated in time-series data*.

Crespo, R., Fotheringham, S., & Charlton, M. (2007). Application of Geographically Weighted Regression to a 19-year set of house price data in London to calibrate local hedonic price models. *Proceedings of the 9th International Conference on GeoComputation*. http://www.geocomputation.org/2007/

De Bruyne, K., & Van Hove, J. (2013). Explaining the spatial variation in housing prices: An economic geography approach. *Applied Economics*, *45*(13), 1673–1689. https://doi.org/10.1080/00036846.2011.636021

Dubin, R. A. (1992). Spatial autocorrelation and neighborhood quality. *Regional Science and Urban Economics*, *22*(3), 433–452. https://doi.org/10.1016/0166-0462(92)90038-3

Ghaderi, J., & Izady, B. (2016). Studying the Effects of Social and Economic Factors on the Housing Prices in Iran (1972-2013). *Urban Economics*, *1*(1), 55–75. https://doi.org/10.22108/ue.2016.22104

Glaeser, E. L., Gyourko, J., & Saiz, A. (2008). Housing supply and housing bubbles. *Journal of Urban Economics*, *64*(2), 198–217. https://doi.org/10.1016/j.jue.2008.07.007

Goodchild, M., & Li, L. (2012). Assuring the Quality of Volunteered Geographic Information. *Spatial Statistics*, *1*, 110–120. https://doi.org/10.1016/j.spasta.2012.03.002

Goodman, A. C., & Thibodeau, T. G. (1998). Housing Market Segmentation. *Journal of Housing Economics*, *7*(2), 121–143. https://doi.org/10.1006/jhec.1998.0229

Gyourko, J., Mayer, C., & Sinai, T. (2013). Superstar Cities. *American Economic Journal: Economic Policy*, *5*(4), 167–199. https://doi.org/10.1257/pol.5.4.167

Habash, A. A., & Unanoglu, M. (2022). Factors influencing prices of residential real estate in turkey. *Florya Chronicles of Political Economy*, *8*(2), Article 2.

Hewitt, C. M., & Hewitt, W. T. (2012). The effect of proximity to urban rail on housing prices in Ottawa. *Journal of Public Transportation*, *15*(4), 43–65.

Hu, S., Yang, S., Li, W., Zhang, C., & Xu, F. (2016). Spatially non-stationary relationships between urban residential land price and impact factors in Wuhan city, China. *Applied Geography*, *68*, 48–56. https://doi.org/10.1016/j.apgeog.2016.01.006

Huang, B., Wu, B., & Barry, M. (2010). Geographically and temporally weighted regression for modeling spatio-temporal variation in house prices. *International Journal of Geographical Information Science*, *24*(3), 383–401. https://doi.org/10.1080/13658810802672469

Kestens, Y., Thériault, M., & Des Rosiers, F. (2006). Heterogeneity in hedonic modelling of house prices: Looking at buyers' household profiles. *Journal of Geographical Systems*, *8*(1), 61–96. https://doi.org/10.1007/s10109-005-0011-8

Kok, S. H., Ismail, N. W., & Lee, C. (2018). The sources of house price changes in Malaysia. *International Journal of Housing Markets and Analysis, 11*(2), 335–355. https://doi.org/10.1108/IJHMA-04-2017-0039

Leishman, C. (2009). Spatial Change and the Structure of Urban Housing Sub-markets. *Housing Studies*, *24*(5), 563–585. https://doi.org/10.1080/02673030903082310

Leung, C. K. Y., Leong, Y. C. F., & Wong, S. K. (2006). Housing Price Dispersion: An Empirical Investigation. *The Journal of Real Estate Finance and Economics*, *32*(3), 357–385. https://doi.org/10.1007/s11146-006-6806-7

Li, S., ye, X., Lee, J., Gong, J., & Qin, C. (2017). Spatiotemporal Analysis of Housing Prices in China: A Big Data Perspective. *Applied Spatial Analysis and Policy*, *10*. https://doi.org/10.1007/s12061-016-9185-3

Liu, N., & Strobl, J. (2023). Impact of neighborhood features on housing resale prices in Zhuhai (China) based on an (M)GWR model. *Big Earth Data*, *7*(1), 146–169. https://doi.org/10.1080/20964471.2022.2031543

Machin, S., & Gibbons, S. (2008). Valuing School Quality, Better Transport, and Lower Crime: Evidence from House Prices. *Oxford Review of Economic Policy*, *24*, 99–119. https://doi.org/10.1093/oxrep/grn008

Mallikharjuna Rao, K., Saikrishna, G., & Supriya, K. (2023). Data preprocessing techniques: Emergence and selection towards machine learning models - a practical review using HPA dataset. *Multimedia Tools and Applications*. https://doi.org/10.1007/s11042-023-15087-5

Manganelli, B., Pontrandolfi, P., Azzato, A., & Murgante, B. (2014). Using geographically weighted regression for housing market segmentation. *International Journal of Business Intelligence and Data Mining*, *9*(2), 161–177. https://doi.org/10.1504/IJBIDM.2014.065100

Mankad, M. D. (2022). Comparing OLS based hedonic model and ANN in house price estimation using relative location. *Spatial Information Research*, *30*(1), 107–116. https://doi.org/10.1007/s41324-021-00416-3

McIlhatton, D., McGreal, W., Taltavul de la Paz, P., & Adair, A. (2016). Impact of crime on spatial analysis of house prices: Evidence from a UK city. *International Journal of Housing Markets and Analysis*, *9*(4), 627–647. https://doi.org/10.1108/IJHMA-10-2015-0065

Mostafa, M. M. (2018). A spatial econometric analysis of residential land prices in Kuwait. *Regional Studies, Regional Science*, *5*(1), 290–311. https://doi.org/10.1080/21681376.2018.1518154

Nistor, A., & Reianu, D. (2018). Determinants of housing prices: Evidence from Ontario cities, 2001-2011. *International Journal of Housing Markets and Analysis*, *11*(3), 541–556. https://doi.org/10.1108/IJHMA-08-2017-0078

Raslanas, S., Tupenaite, L., & Đteinbergas, T. (2006). Research on the prices of flats in the South East London and Vilnius. *International Journal of Strategic Property Management*, *10*. https://doi.org/10.1080/1648715X.2006.9637544

Saiz, A. (2010). The Geographic Determinants of Housing Supply. *The Quarterly Journal of Economics*, *125*(3), 1253–1296.

Schmidt, A. F., & Finan, C. (2018). Linear regression and the normality assumption. *Journal of Clinical Epidemiology*, *98*, 146–151. https://doi.org/10.1016/j.jclinepi.2017.12.006

Singh, A. (2024, January 3). Decoding the Power of Proximity: A Beginner's Guide to K-Nearest Neighbors (KNN) in Machine…. *Medium*. https://medium.com/@ashpalsingh0909/decoding-the-power-of-proximity-a-beginners-guide-to-k-nearest-neighbors-knn-in-machine-16b9f06d1ac5

Souza, T. G. D., Fonseca, F. D. R., Fernandes, V. D. O., & Pedrassoli, J. C. (2021). Exploring spatial analysis of housing prices obtained from Web Scraping Technique.

*The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLIII-B4-2021*, 135–140. XXIV ISPRS Congress <q>Imaging today, foreseeing tomorrow</q>, Commission IV - 2021 edition, 5–9 July 2021. https://doi.org/10.5194/isprs-archives-XLIII-B4-2021-135-2021

Sun, H., Tu, Y., & Yu, S.-M. (2005). A Spatio-Temporal Autoregressive Model for Multi-Unit Residential Market Analysis*. *The Journal of Real Estate Finance and Economics*, *31*(2), 155–187. https://doi.org/10.1007/s11146-005-1370-0

Tomal, M., & Helbich, M. (2023). A spatial autoregressive geographically weighted quantile regression to explore housing rent determinants in Amsterdam and Warsaw. *Environment and Planning B: Urban Analytics and City Science*, *50*(3), 579–599. https://doi.org/10.1177/23998083221122790

Wang, W.-C., Chang, Y.-J., & Wang, H.-C. (2019). An Application of the Spatial Autocorrelation Method on the Change of Real Estate Prices in Taitung City. *ISPRS International Journal of Geo-Information*, *8*(6), Article 6. https://doi.org/10.3390/ijgi8060249

Zhao, H., Ren, Z., & Tan, J. (2018). The Spatial Patterns of Land Surface Temperature and Its Impact Factors: Spatial Non-Stationarity and Scale Effects Based on a Geographically-Weighted Regression Model. *Sustainability*, *10*(7), Article 7. https://doi.org/10.3390/su10072242

## ONLINE REFERENCES

URL1, World Population Prospects 2019 Highlights. (n.d.). Retrieved January 17, 2024, from World Population Prospects 2019 Highlights (un.org)

URL2, TÜİK Kurumsal. (n.d.). Retrieved February 24, 2024, from TurkStat Corporate (tuik.gov.tr)

URL3, Specialist_Put_4800. (2021, October 19). Retrieved December 15, 2023, from Population change in Turkish provinces during 2000 and 2020 : r/Turkey (reddit.com)

URL4, Turkish property prices are soaring. (n.d.). Retrieved December 3, 2023, from Turkish property prices are soaring (economist.com)

URL5, How OLS regression works—ArcGIS Pro | Documentation. (n.d.). Retrieved May 25, 2023, from How OLS regression works—ArcGIS Pro | Documentation

URL6, Columbia University Mailman School of Public Health. (2016). Retrieved May 24, 2023, from Geographically Weighted Regression | Columbia University Mailman School of Public Health

URL7, Ankara, Turkey Metro Area Population 1950-2023. (n.d.). Retrieved May 24, 2023, from Ankara, Turkey Metro Area Population 1950-2024 | MacroTrends

URL8, sahibinden.com—Türkiye'nin Online İlan ve Alışveriş Sitesi. (n.d.). Retrieved May 25, 2023, from sahibinden.com - Türkiye'nin Online İlan ve Alışveriş Sitesi

URL9, Başarsoft: Coğrafi Bilgi Sistemleri (CBS-GIS) ve Altyapı Çözümleri. (n.d.). Retrieved January 23, 2024, from Başarsoft: Coğrafi Bilgi Sistemleri (CBS-GIS) ve Altyapı Çözümleri (basarsoft.com.tr)

URL10, How Spatial Autocorrelation (Global Moran's I) works—ArcGIS Pro | Documentation. (n.d.). Retrieved June 7, 2023, from How Spatial Autocorrelation (Global Moran's I) works—ArcGIS Pro | Documentation

URL11, Kriging (Spatial Analyst)—ArcGIS Pro | Documentation. (n.d.). Retrieved December 3, 2023, from Kriging (Spatial Analyst)—ArcGIS Pro | Documentation

URL12, Understanding a semivariogram: The range, sill, and nugget—ArcGIS Pro | Documentation. (n.d.). Retrieved January 23, 2024, from Understanding a semivariogram: The range, sill, and nugget—ArcGIS Pro | Documentation

URL13, (Ordinary Least Squares (OLS) (Spatial Statistics)—ArcGIS Pro | Documentation, n.d.). Retrieved May 27, 2023, from Ordinary Least Squares (OLS) (Spatial Statistics)—ArcGIS Pro | Documentation

URL14, *TCMB*. (n.d.). Retrieved October 5, 2024, Retrieved October 5, 2024, from from CBT (tcmb.gov.tr)

URL15, *ankara cayyolu map—Google Search*. (n.d.). Retrieved October 7, 2024, from ankara kizilay map - Google Search

URL16, *ankara kizilay map—Google Search*. (n.d.). Retrieved October 7, 2024, from ankara cayyolu map - Google Search

URL17, *Ankara Trafik Haritası | Ankara Yol Durumu*. (n.d.). Retrieved October 7, 2024, from Ankara Trafik Haritası | Ankara Yol Durumu (hurriyet.com.tr)

URL18, *OpenStreetMap. (n.d.). OpenStreetMap. Retrieved October 28, 2024,* from OpenStreetMap

# APPENDICES

## A. Variables Characteristics.

Table A.1: Variables Characteristics.

| Index | Variable | Description in Turkish | Description in English |
|---|---|---|---|
| **V1** | Price | TL cinsinden ev fiyatı. | Housing price in TL. |
| **V2** | Area Per m2 | Metrekare başına ev alanı. | Housing area per square meter. |
| **V3** | No. of Rooms | Evdeki toplam oda sayısı. | Total number of rooms in the house. |
| **V4** | Floor No. | Ev kat numarası. | House floor number. |
| **V5** | dist. Transport | Binaların en yakın ulaşım istasyonuna metre cinsinden uzaklığı. | Distance of the buildings in meters to the closest transportation station. |
| **V6** | dist. University | Binaların en yakın üniversiteye metre cinsinden uzaklığı. | Distance of between buildings and the closest university in meters. |
| **V7** | dist. Mall | Binaların en yakın AVM'ye metre cinsinden uzaklığı. | Distance of the buildings in meters to the closest shopping mall. |
| **V8** | dist. Hospital | Binaların en yakın hastaneye metre cinsinden uzaklığı. | Distance of the buildings in meters to the closest hospital. |
| **V9** | dist. School | Binaların en yakın okula metre cinsinden uzaklığı. | Distance of the buildings in meters to the closest school. |
| **V10** | dist. Bank | Binaların en yakın bankaya metre cinsinden uzaklığı. | Distance of the buildings in meters to the closest bank. |
| **V11** | dist. ATM | Binaların en yakın ATM'ye metre cinsinden uzaklığı. | Distance of the buildings in meters to the closest ATM. |
| **V12** | dist. Supermarket | Binaların en yakın süpermarkete metre cinsinden uzaklığı. | Distance of the buildings in meters to the closest Supermarket. |
| **V13** | dist. Budget Supermarket | Binaların en yakın bütçe süpermarketine metre cinsinden uzaklığı. | Distance of the buildings in meters to the closest budget supermarket. |

## B. Variables Histograms Including Outliers.

(a)

(b)

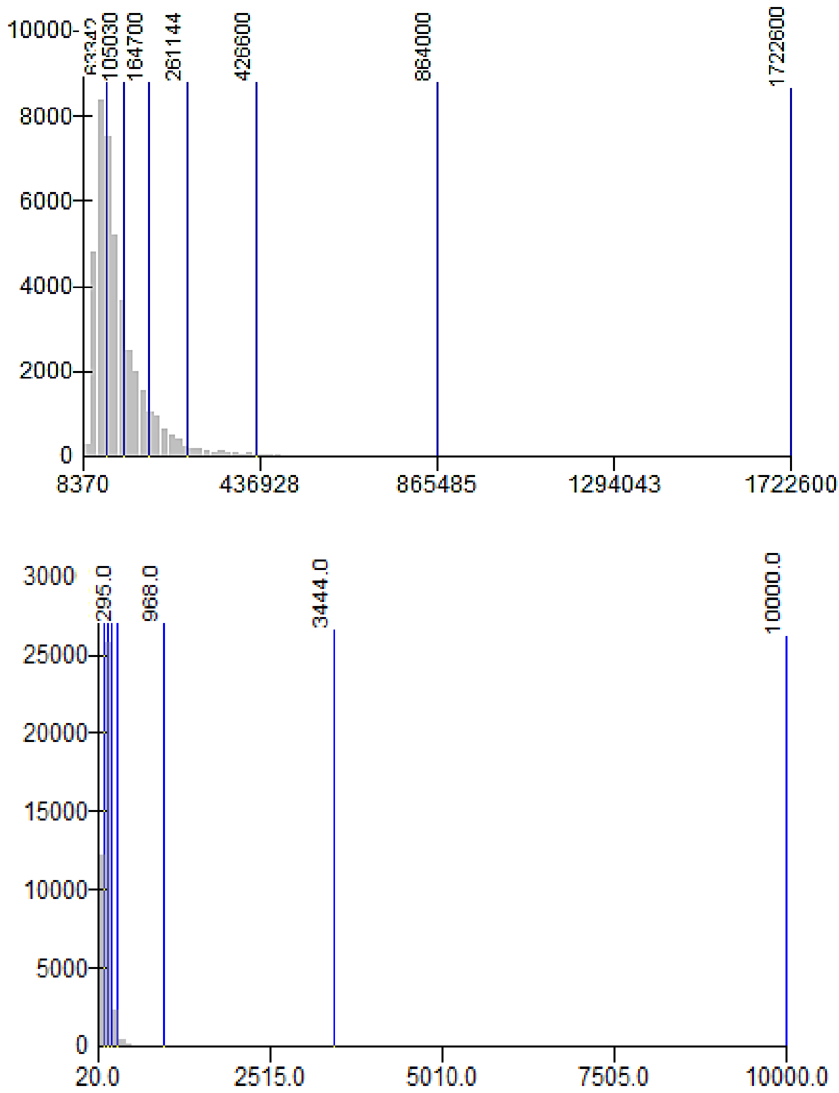



Figure B.1: Histograms of a) Housing prices in in (USD)  b) Area per m2.

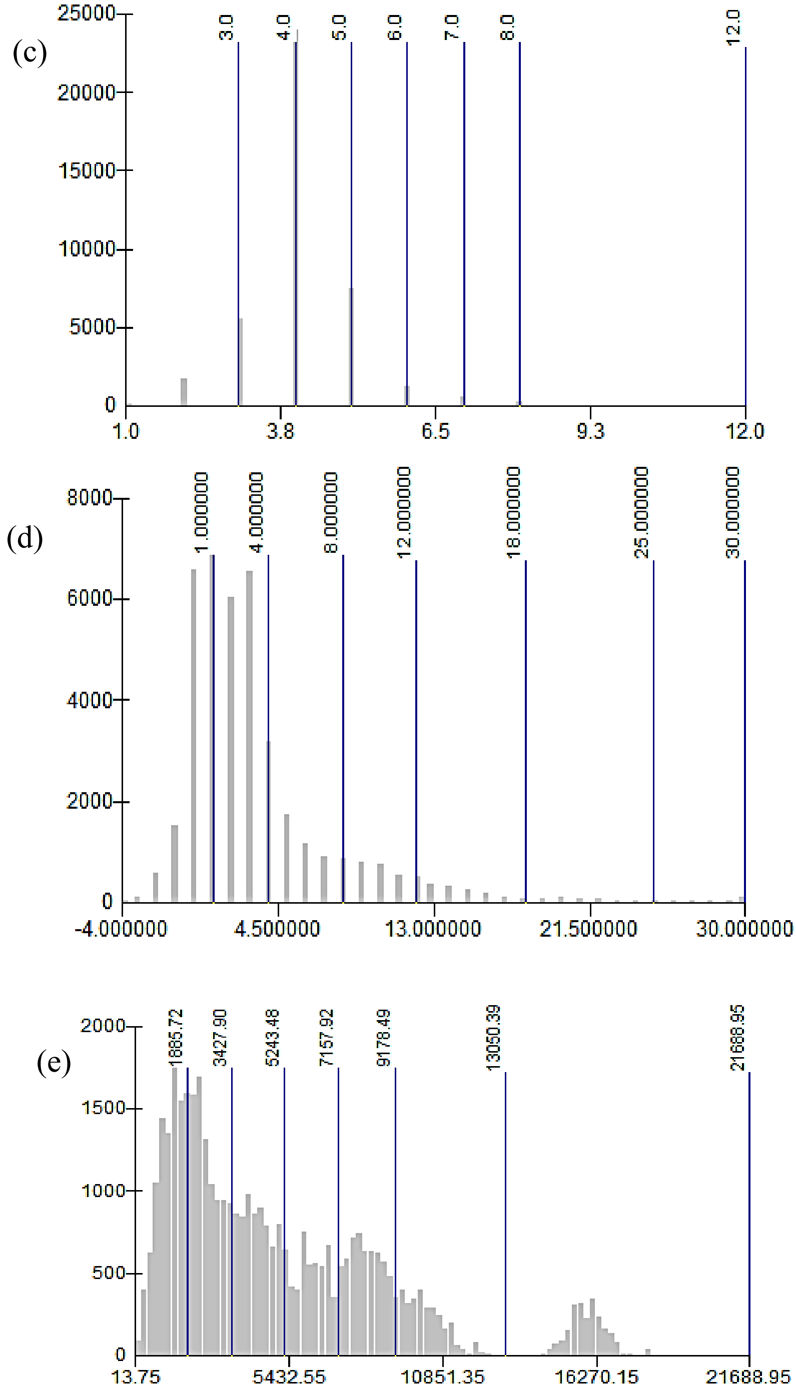



Figure B.2: Histograms of c) Number of rooms d) Floor number e) Nearest distance to universities.

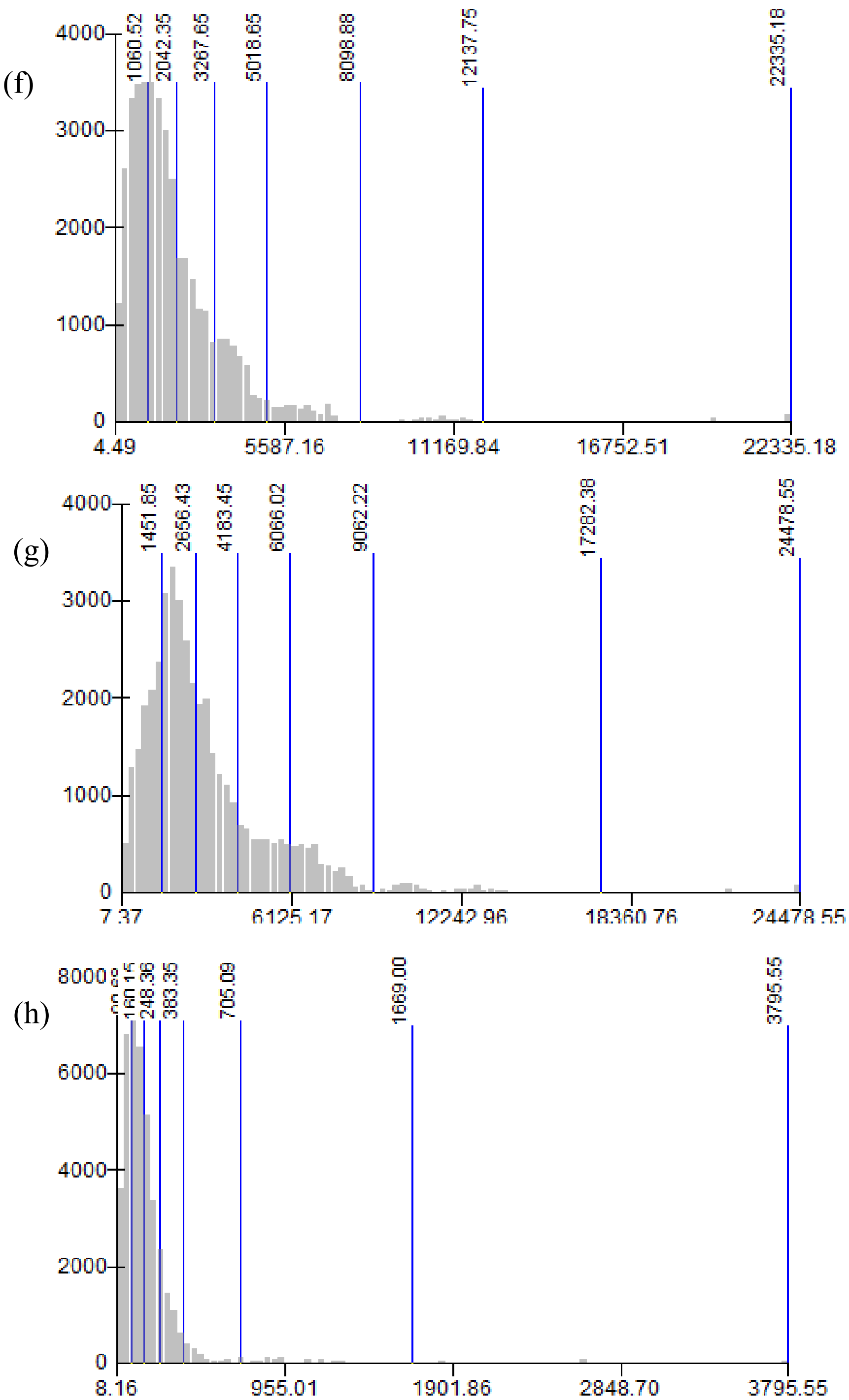



Figure B.3: Histogram of f) Nearest distance to hospitals g) Nearest distance to malls h) Nearest distance to transportation.

(i)

(j)

(k)

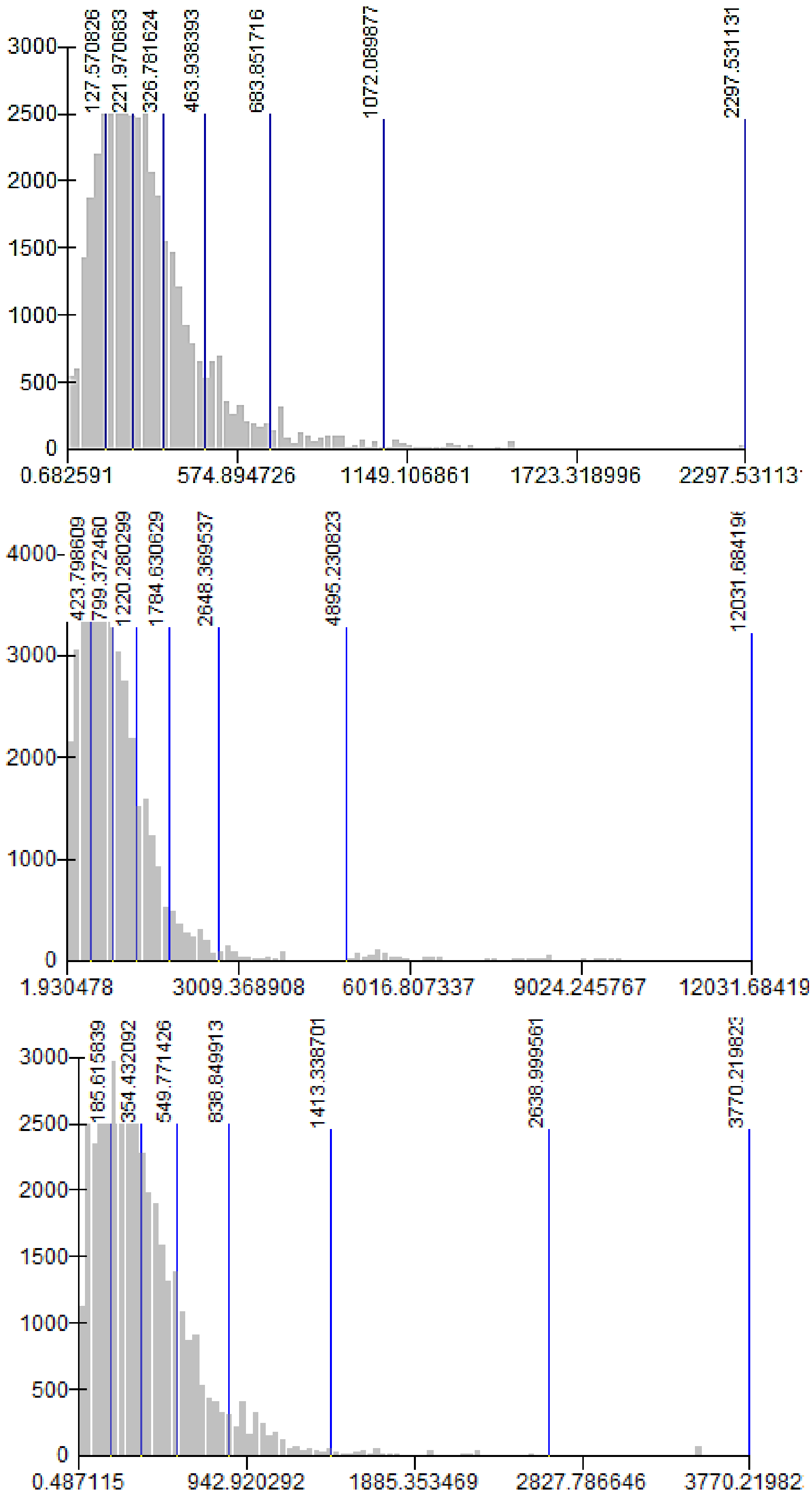
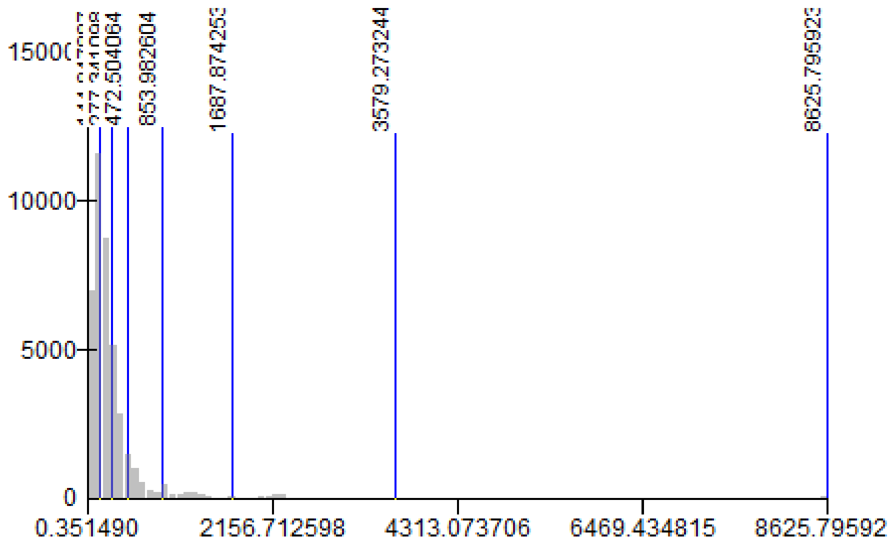



Figure B.4: Histogram of i) Nearest distance to schools j) Nearest distance to Bank k) Nearest distance to ATM.

(l)




(m)

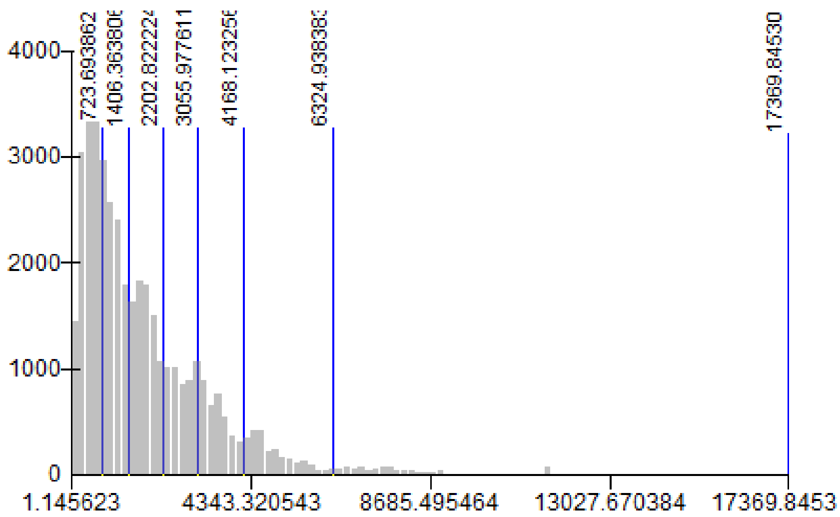



Figure B.5: Histogram of l) Nearest distance  to Budget supermarkets (A101 & BIM) m) Nearest distance  to Supermaarkets (Migros).

## C. Customized GIS Application Code in C#

### Form 1: Main interface

```
using System;
using System.Collections.Generic;
using System.ComponentModel;
using System.Data;
using System.Drawing;
using System.Linq;
using System.Text;
using System.Threading.Tasks;
using System.Windows.Forms;
using MapInfo;
using System.Runtime.InteropServices;
using System.Diagnostics;




namespace Spatial_Analysis_On_Housing_Prices_Ankara
{
    public partial class Form1 : Form
    {
        public static MapInfo.MapInfoApplication mi;

        public static string win_id;
        [DllImport("user32.dll")]
        static extern bool MoveWindow(IntPtr hWnd, int X, int Y, int
nWidth, int nHeight, bool bRepaint);

        CallBack callb;
        public Ankara_Neighbourhood_Histogram f3 = new
Ankara_Neighbourhood_Histogram();

        public Form1()
        {
            InitializeComponent();
            callb = new CallBack(this);

        }

        private void Form1_Load(object sender, EventArgs e)
        {

            Process[] _proceses = null;
            _proceses = Process.GetProcessesByName("MapInfo");
            foreach (Process proces in _proceses)
            {
                proces.Kill();
            }

            mi = new MapInfo.MapInfoApplication();
```

```
            int p = Ankara_Map.Handle.ToInt32();
            mi.Do("set next document parent " + p.ToString() + "style 
1");
            mi.Do("set application window " + p.ToString());
            mi.Do("run application \"" + "C:/Users/Toshiba/Desktop/GGIT-
1ST/3rd_semester/Thesis/Work/MapInfo/Spatial_Analysis_on_Housing_Prices_
Ankara_City.WOR" + "\"");
            mi.SetCallback(callb);
            mi.Do("create buttonpad \"a\" as toolbutton calling OLE 
\"info\" id 2001");
            win_id = mi.Eval("frontwindow()");

        }



        private void neighboursToolStripMenuItem_Click(object sender, 
EventArgs e)
        {
            Ankara_Neighborhood_Thematic f2 = new 
Ankara_Neighborhood_Thematic();
            f2.Show();
        }

        private void Info_Click(object sender, EventArgs e)
        {
            mi.Do("run menu command id 2001");

        }

        private void Form1_Resize(object sender, EventArgs e)
        {
            if (mi != null)
            {
                // The form has been resized.
                if (mi.Eval("WindowID(0)") != "")
                {
                    // Update the map to match the current size of the 
panel.

MoveWindow((System.IntPtr)long.Parse(mi.Eval("WindowInfo(FrontWindow(),1
2)")), 0, 0, this.Ankara_Map.Width, this.Ankara_Map.Height, false);
                }
            }

        }

        private void button1_Click(object sender, EventArgs e)
        {
            Sorted_Neighbourhoods_Houses_Pricing f5 = new 
Sorted_Neighbourhoods_Houses_Pricing();
            f5.Show();
        }

        private void button2_Click(object sender, EventArgs e)
        {
            mi.Do("run menu command  1705");
```

```
        }

        private void button3_Click(object sender, EventArgs e)
        {
            mi.Do("run menu command  1706");
        }

        private void button4_Click(object sender, EventArgs e)
        {
            mi.Do("run menu command 1702");
        }

        private void buttonCreateObjectAndSelectData_Click_Click(object
sender, EventArgs e)
        {
            Ankara_Neighbourhood_Histogram histogramForm = new
Ankara_Neighbourhood_Histogram();
            histogramForm.CreateObjectAndSelectData();
        }
    }
}
```

## Ankara_Neighborhood_Thematic Form:

```
using System;
using System.Collections.Generic;
using System.ComponentModel;
using System.Data;
using System.Drawing;
using System.Linq;
using System.Text;
using System.Threading.Tasks;
using System.Windows.Forms;
using MapInfo;

namespace Spatial_Analysis_On_Housing_Prices_Ankara
{
    public partial class Ankara_Neighborhood_Thematic : Form
    {
        public Ankara_Neighborhood_Thematic()
        {
            InitializeComponent();
            PopulateComboBox();
            PopulateListBox();
            comboBoxDistricts.SelectedIndexChanged += new
EventHandler(comboBoxDistricts_SelectedIndexChanged);
            listBox1.SelectedIndexChanged += new
EventHandler(listBox1_SelectedIndexChanged);
        }

        private void PopulateComboBox()
```

```
        {
            // Add district names to the ComboBox
            comboBoxDistricts.Items.Add("Çankaya");
            comboBoxDistricts.Items.Add("Golbasi");
            comboBoxDistricts.Items.Add("Mamak");
            comboBoxDistricts.Items.Add("Yenimahalle");
            comboBoxDistricts.Items.Add("Kecioren");
            comboBoxDistricts.Items.Add("Altindag");
            comboBoxDistricts.Items.Add("Sincan");
            comboBoxDistricts.Items.Add("Etimesgut");
            comboBoxDistricts.Items.Add("Ankare Metropolitan Area
Population");

        }

        private void PopulateListBox()
        {
            // Add thematic categories to the ListBox

            listBox1.Items.Add("population_2021");
            listBox1.Items.Add("population_2020");
            listBox1.Items.Add("population_2019");
            listBox1.Items.Add("population_2018");
            listBox1.Items.Add("population_2017");
            listBox1.Items.Add("population_2016");
            listBox1.Items.Add("population_2015");
            listBox1.Items.Add("population_2014");


        }

        private void comboBoxDistricts_SelectedIndexChanged(object
sender, EventArgs e)
        {
            ApplySelectedThematicMapping();
        }

        private void listBox1_SelectedIndexChanged(object sender,
EventArgs e)
        {
            ApplySelectedThematicMapping();
        }

        private void ApplySelectedThematicMapping()
        {
            if (comboBoxDistricts.SelectedItem == null ||
listBox1.SelectedItem == null)
                return;

            string selectedDistrict =
comboBoxDistricts.SelectedItem.ToString();
            string selectedCategory = listBox1.SelectedItem.ToString();

            if (selectedCategory.StartsWith("population"))
            {
                ApplyPopulationThematicMapping(selectedCategory);
            }
```

```
            else
            {
                ApplyThematicMapping(selectedDistrict,
selectedCategory);
            }
        }

        private void ApplyThematicMapping(string district, string
thematic_column)
        {
            int n = Convert.ToInt16(textBox1.Text);
            string p = panel1.Handle.ToString();
            removetematik();

            // Select the data for the selected district
            Form1.mi.Do("select * from thiessen_all where Ilcedi=\"" +
district + "\" into thiessen_sel");
            Form1.mi.Do("add map layer thiessen_sel");
            Form1.mi.Do("select Max(" + thematic_column + ") from
thiessen_sel into maxselect");
            double maxx =
Convert.ToDouble(Form1.mi.Eval("maxselect.col1"));
            Form1.mi.Do("select Min(" + thematic_column + ") from
thiessen_sel into minselect");
            double minn =
Convert.ToDouble(Form1.mi.Eval("minselect.col1"));

            Form1.mi.Do("select " + thematic_column + " from
thiessen_sel order by " + thematic_column + " into sel noselect");

            int diff = Convert.ToInt32(maxx - minn);
            int range = diff / n;
            int c_range = Convert.ToInt16(255 / n);

            Form1.mi.Do("fetch first from sel");
            string r1 = Convert.ToString(Form1.mi.Eval("sel.col1"));
            string r2 = string.Empty;
            string cmstr = string.Empty;

            for (int i = 1; i < n; i++)
            {
                int temp = Convert.ToInt32(r1) + range;
                r2 = temp.ToString();
                string rgb = Convert.ToString(Form1.mi.Eval("RGB(255," +
((n - i) * c_range).ToString() + "," + ((n - i) * c_range).ToString() +
")"));
                cmstr = cmstr + r1 + ":" + r2 + " brush(2," + rgb +
",16777215) pen(1,1,0), ";
                r1 = r2;
            }

            Form1.mi.Do("fetch last from sel");
            r2 = Convert.ToString(Form1.mi.Eval("sel.col1"));
            cmstr = cmstr + r1 + ":" + r2 + " brush(2,16711680,16777215)
pen(1,1,0)";
```

```
            Form1.mi.Do("shade window " + Form1.win_id.ToString() + "
thiessen_sel with " + thematic_column + " ranges apply all use color
Brush (2,16711680,16777215) " + cmstr);
            Form1.mi.Do("Set Next Document Parent " + p + " Style 1");
            Form1.mi.Do("Create Cartographic Legend From Window " +
Form1.win_id.ToString() + " Behind Frame From Layer 1");
        }

        private void ApplyPopulationThematicMapping(string
thematic_column)
        {
            int n = Convert.ToInt16(textBox1.Text);
            string p = panel1.Handle.ToString();
            removetematik();

            // Select the data from the Metropolitan_Neighbourhoods
layer

            Form1.mi.Do("select Max(" + thematic_column + ") from
Metropolitan_Neighbourhoods into maxselect");
            double maxx =
Convert.ToDouble(Form1.mi.Eval("maxselect.col1"));
            Form1.mi.Do("select Min(" + thematic_column + ") from
Metropolitan_Neighbourhoods into minselect");
            double minn =
Convert.ToDouble(Form1.mi.Eval("minselect.col1"));

            Form1.mi.Do("select " + thematic_column + " from
Metropolitan_Neighbourhoods order by " + thematic_column + " into sel
noselect");

            int diff = Convert.ToInt32(maxx - minn);
            int range = diff / n;
            int c_range = Convert.ToInt16(255 / n);

            Form1.mi.Do("fetch first from sel");
            string r1 = Convert.ToString(Form1.mi.Eval("sel.col1"));
            string r2 = string.Empty;
            string cmstr = string.Empty;

            for (int i = 1; i < n; i++)
            {
                int temp = Convert.ToInt32(r1) + range;
                r2 = temp.ToString();
                string rgb = Convert.ToString(Form1.mi.Eval("RGB(255," +
((n - i) * c_range).ToString() + "," + ((n - i) * c_range).ToString() +
")"));
                cmstr = cmstr + r1 + ":" + r2 + " brush(2," + rgb +
",16777215), ";
                r1 = r2;
            }

            Form1.mi.Do("fetch last from sel");
            r2 = Convert.ToString(Form1.mi.Eval("sel.col1"));
            cmstr = cmstr + r1 + ":" + r2 + "
brush(2,16711680,16777215)";
```

```
            Form1.mi.Do("shade window " + Form1.win_id.ToString() + "
Metropolitan_Neighbourhoods with " + thematic_column + " ranges apply
all use color Brush (2,16711680,16777215) " + cmstr);
            Form1.mi.Do("Set Next Document Parent " + p + " Style 1");
            Form1.mi.Do("Create Cartographic Legend From Window " +
Form1.win_id.ToString() + " Behind Frame From Layer 11");
            //Form1.mi.Do("Set Map Window " + Form1.win_id.ToString() +
" Order Layers 4 DestGroupLayer 0 Position 2");
        }


        public void removetematik()
        {
            for (int k = Convert.ToInt16(Form1.mi.Eval("mapperinfo(" +
Form1.win_id + ",9)")); k > 0; k = k - 1)
            {
                int value = Convert.ToInt16(Form1.mi.Eval("layerinfo(" +
Form1.win_id + "," + Convert.ToString(k) + ",24)"));

                if (value == 3)
                {
                    Form1.mi.Do("remove map layer \"" +
Form1.mi.Eval("layerinfo(" + Form1.win_id + "," + Convert.ToString(k) +
",1)") + "\"");
                }
            }
        }

        private void Ankara_Neighborhood_Thematic_Load(object sender,
EventArgs e)
        {
            listBox1.SelectedIndex = 0;
        }


        private void button2_Click(object sender, EventArgs e)
        {
            Form1.mi.Do("Select district_name, sum(population_2021) from
Metropolitan_Neighbourhoods group by district_name order by
population_2021 desc into sel noselect ");

            int range =
Convert.ToInt16(Form1.mi.Eval("int(tableinfo(sel,8)/" +
Convert.ToString("9") + ")"));

            Form1.mi.Do("fetch first from sel");

            string r1 = null; //
Convert.ToString(Form1.mi.Eval("sel.col1"));
            string r2 = null;
//Convert.ToString(Form1.mi.Eval("sel.col2"));

            int Cvalues = (range * 9) + 1;

            //listBox1.Items.Clear();

            for (int i = 1; i < Cvalues; i++)
```

```
            {
                Form1.mi.Do("fetch rec " + Convert.ToString(i) + " from 
sel");
                r1 = Convert.ToString(Form1.mi.Eval("sel.col1"));
                r2 = Convert.ToString(Form1.mi.Eval("sel.col2"));


                string[] row = new string[] { r1, r2 };
                dataGridView1.Rows.Add(row);

                //listBox1.Items.Add(r1 + "   " + r2+"   "+r3);
            }

        }

    }
}
```

**Ankara_Neighbourhood_Histogram Form:**

```
using System;

using System.Collections.Generic;

using System.ComponentModel;

using System.Data;

using System.Drawing;

using System.Linq;

using System.Text;

using System.Threading.Tasks;

using System.Windows.Forms;

using System.Runtime.InteropServices;



namespace Spatial_Analysis_On_Housing_Prices_Ankara

{
```

```
    public partial class Ankara_Neighbourhood_Histogram : Form
    {
        public static MapInfo.MapInfoApplication mi;
        [DllImport("user32.dll")]
        static extern bool MoveWindow(IntPtr hWnd, int X, int Y, int nWidth, int
nHeight, bool bRepaint);
        private Dictionary<string, string> textMappings = new Dictionary<string,
string>
        {
            { "Çankaya",
                "1: 485000 -  1282500 " +
                "2: 1282501 - 1787500 " +
                "3: 1787501 - 2344333 " +
                "4: 2344334 - 3080000 " +
                "5: 3080001 - 4050000" },
            { "Altindag",
                "1: 325000 -  1040000 " +
                "2: 1040000 - 1755000 " +
                "3: 1755000 - 2470000 " +
                "4: 2470000 - 3185000 " +
                "5: 3185000 - 3900000" },
            { "Mamak",
                "1: 395000 -  1106000 " +
```

```
    "2: 1106000 - 1817000 " +
    "3: 1817000 - 2528000 " +
    "4: 2528000 - 3239000 " +
    "5: 3239000 - 3950000" },
{ "Kecioren",
    "1: 425000 -  1139999 " +
    "2: 1139999 - 1854998 " +
    "3: 1854998 - 2569997 " +
    "4: 2569997 - 3284996 " +
    "5: 3284996 - 3999999" },
{ "Yenimahalle",
    "1: 325000 -  1060000 " +
    "2: 1060000 - 1795000 " +
    "3: 1795000 - 2530000 " +
    "4: 2530000 - 3265000 " +
    "5: 3265000 - 4000000" },
{ "Etimesgut",
    "1: 379000 -  1103000 " +
    "2: 1103000 - 1827000 " +
    "3: 1827000 - 2551000 " +
    "4: 2551000 - 3275000 " +
    "5: 3275000 - 3999000" },
```

```
    { "Sincan",
        "1: 389000 -  937200 " +
        "2: 937200 -  1485400 " +
        "3: 1485400 - 2033600 " +
        "4: 2033600 - 2581800 " +
        "5: 2581800 - 3130000" },
    { "Golbasi",
        "1: 560000 -  1214666 " +
        "2: 1214666 - 1869332 " +
        "3: 1869332 - 2523998 " +
        "4: 2523998 - 3178664 " +
        "5: 3178664 - 3833333" },

    { "Pursaklar",
        "1: 410000 -  971800 " +
        "2: 971800 -  1533600 " +
        "3: 1533600 - 2095400 " +
        "4: 2095400 - 2657200 " +
        "5: 2657200 - 3219000" },
};
public Ankara_Neighbourhood_Histogram()
{
```

```
        InitializeComponent();
    }
    public void fill_form()
    {
        label2.Text = Form1.mi.Eval("Thiessen_all.Ilcedi");
        //Form1.mi.Do("select * from Thiessen_all where Name=\"" + label2.Text + "\" into Sel1 noselect");
        UpdateTextBox();
        // Highlight the selected province boundaries
        Form1.mi.Do("select * from Ankara_Metropolitan_Area where Ilceadi=\"" + label2.Text + "\"  into Sel_ILCE noselect");
        Form1.mi.Do("add map layer sel_ILCE");
        Form1.mi.Do("Set Map Window frontwindow() layer sel_ILCE display Global Global Pen (3,66,16711680) Global Brush (1,16777215,16777215)");
        Form1.mi.Do("select range_id, sum(No_Mall) from Thiessen_all where Ilcedi=\"" + label2.Text + "\" group by range_id order by range_id into Sel1 noselect");

        int p1 = tabPage4.Handle.ToInt32();
            Form1.mi.Do("set next document parent " + p1.ToString() + "style 1");
            Form1.mi.Do("Graph Range_id, COL2  From sel1 Using \"" + "C:/Users/Toshiba/Desktop/GGIT-1ST/3rd_semester/Thesis/Work/MapInfo/Graphs/No_Mall.3tf" + "\"Series In Columns");
```

```
Form1.mi.Do("set graph title \"No_Mall\"");

Form1.mi.Do("select range_id, sum(No_School) from Thiessen_all
where Ilcedi=\"" + label2.Text + "\" group by range_id order by range_id into Sel2
noselect");

int p2 = tabPage1.Handle.ToInt32();

Form1.mi.Do("set next document parent " + p2.ToString() + "style 1");

Form1.mi.Do("Graph Range_id, COL2  From sel2 Using \"" +
"C:/Users/Toshiba/Desktop/GGIT-
1ST/3rd_semester/Thesis/Work/MapInfo/Graphs/No_Mall.3tf" + "\"Series In
Columns");

Form1.mi.Do("set graph title \"No_School\"");

Form1.mi.Do("select range_id, sum(No_University) from Thiessen_all
where Ilcedi=\"" + label2.Text + "\" group by range_id order by range_id into Sel3
noselect");

int p3 = tabPage2.Handle.ToInt32();

Form1.mi.Do("set next document parent " + p3.ToString() + "style 1");

Form1.mi.Do("Graph Range_id, COL2  From sel3 Using \"" +
"C:/Users/Toshiba/Desktop/GGIT-
1ST/3rd_semester/Thesis/Work/MapInfo/Graphs/No_Mall.3tf" + "\"Series In
Columns");

Form1.mi.Do("set graph title \"No_University\"");

Form1.mi.Do("select range_id, sum(No_Hospital) from Thiessen_all
where Ilcedi=\"" + label2.Text + "\" group by range_id order by range_id into Sel4
noselect");

int p4 = tabPage3.Handle.ToInt32();
```

```
Form1.mi.Do("set next document parent " + p4.ToString() + "style 1");
Form1.mi.Do("Graph Range_id, COL2  From sel4 Using \"" +
"C:/Users/Toshiba/Desktop/GGIT-
1ST/3rd_semester/Thesis/Work/MapInfo/Graphs/No_Mall.3tf" + "\"Series In
Columns");
Form1.mi.Do("set graph title \"No_Hospital\"");
Form1.mi.Do("select range_id, sum(No_ATM) from Thiessen_all where
Ilcedi=\"" + label2.Text + "\" group by range_id order by range_id into Sel5
noselect");
int p5 = tabPage5.Handle.ToInt32();
Form1.mi.Do("set next document parent " + p5.ToString() + "style 1");
Form1.mi.Do("Graph Range_id, COL2  From sel5 Using \"" +
"C:/Users/Toshiba/Desktop/GGIT-
1ST/3rd_semester/Thesis/Work/MapInfo/Graphs/No_Mall.3tf" + "\"Series In
Columns");
Form1.mi.Do("set graph title \"No_ATM\"");
Form1.mi.Do("select range_id, sum(No_Bank) from Thiessen_all where
Ilcedi=\"" + label2.Text + "\" group by range_id order by range_id into Sel6
noselect");
int p6 = tabPage6.Handle.ToInt32();
Form1.mi.Do("set next document parent " + p6.ToString() + "style 1");
Form1.mi.Do("Graph Range_id, COL2  From sel6 Using \"" +
"C:/Users/Toshiba/Desktop/GGIT-
1ST/3rd_semester/Thesis/Work/MapInfo/Graphs/No_Mall.3tf" + "\"Series In
Columns");
Form1.mi.Do("set graph title \"No_Bank\"");
```

```
Form1.mi.Do("select range_id, sum(No_Budget_Supermarket) from Thiessen_all where Ilcedi=\"" + label2.Text + "\" group by range_id order by range_id into Sel7 noselect");

int p7 = tabPage7.Handle.ToInt32();

Form1.mi.Do("set next document parent " + p7.ToString() + "style 1");

Form1.mi.Do("Graph Range_id, COL2  From sel7 Using \"" + "C:/Users/Toshiba/Desktop/GGIT-1ST/3rd_semester/Thesis/Work/MapInfo/Graphs/No_Mall.3tf" + "\"Series In Columns");

Form1.mi.Do("set graph title \"No_Budget_Supermarket\"");

Form1.mi.Do("select range_id, sum(No_Migros) from Thiessen_all where Ilcedi=\"" + label2.Text + "\" group by range_id order by range_id into Sel8 noselect");

int p8 = tabPage8.Handle.ToInt32();

Form1.mi.Do("set next document parent " + p8.ToString() + "style 1");

Form1.mi.Do("Graph Range_id, COL2  From sel8 Using \"" + "C:/Users/Toshiba/Desktop/GGIT-1ST/3rd_semester/Thesis/Work/MapInfo/Graphs/No_Mall.3tf" + "\"Series In Columns");

Form1.mi.Do("set graph title \"No_Supermarket\"");

Form1.mi.Do("select range_id, sum(No_Transportation) from Thiessen_all where Ilcedi=\"" + label2.Text + "\" group by range_id order by range_id into Sel9 noselect");

int p9 = tabPage9.Handle.ToInt32();
```

```
            Form1.mi.Do("set next document parent " + p9.ToString() + "style 1");
            Form1.mi.Do("Graph Range_id, COL2  From sel9 Using \"" +
"C:/Users/Toshiba/Desktop/GGIT-
1ST/3rd_semester/Thesis/Work/MapInfo/Graphs/No_Mall.3tf" + "\"Series In
Columns");
            Form1.mi.Do("set graph title \"No_Transportation\"");
         this.ShowDialog();
      }
      private void UpdateTextBox()
      {
         if (textMappings.TryGetValue(label2.Text, out string newText))
         {
            textBox1.Text = newText;
         }
         else
         {
            textBox1.Text = "Default text if no mapping is found"; // Optional
default text
         }
      }
      public void CreateObjectAndSelectData()
      {
         label2.Visible = false;
```

```
label1.Visible = false;
Form1.mi.Do("dim o as object");
Form1.mi.Do("create object as merge from selection into variable o");
// Add your additional values for the case when multiple provinces are selected
string multiProvinceText =
    "1:  320000 -  840000 " +
    "2:  840000 - 1170000 " +
    "3:  1170000 - 1640000 " +
    "4:  1640000 - 2390000 " +
    "5:  2390000 - 4050000";
// Update textBox1 with the additional values for multiple provinces selection
textBox1.Text = multiProvinceText;

Form1.mi.Do("select Range_id_all_provincies, sum(No_Mall) from thiessen_all where obj intersects o group by Range_id_all_provincies order by Range_id_all_provincies into sel");
int p1 = tabPage4.Handle.ToInt32();
Form1.mi.Do("set next document parent " + p1.ToString() + " style 1");
Form1.mi.Do("Graph Range_id_all_provincies, COL2 From sel Using \"" + "C:/Users/Toshiba/Desktop/GGIT-1ST/3rd_semester/Thesis/Work/MapInfo/Graphs/No_Mall.3tf" + "\"Series In Columns");
```

Form1.mi.Do("set graph title \"No_Mall\"");

Form1.mi.Do("select Range_id_all_provincies, sum(No_School) from thiessen_all where obj intersects o group by Range_id_all_provincies order by Range_id_all_provincies into sel2");

int p2 = tabPage1.Handle.ToInt32();

Form1.mi.Do("set next document parent " + p2.ToString() + " style 1");

Form1.mi.Do("Graph Range_id_all_provincies, COL2 From sel2 Using \"" + "C:/Users/Toshiba/Desktop/GGIT-1ST/3rd_semester/Thesis/Work/MapInfo/Graphs/No_Mall.3tf" + "\"Series In Columns");

Form1.mi.Do("set graph title \"No_School\"");

Form1.mi.Do("select Range_id_all_provincies, sum(No_University) from thiessen_all where obj intersects o group by Range_id_all_provincies order by Range_id_all_provincies into sel3");

int p3 = tabPage2.Handle.ToInt32();

Form1.mi.Do("set next document parent " + p3.ToString() + " style 1");

Form1.mi.Do("Graph Range_id_all_provincies, COL2 From sel3 Using \"" + "C:/Users/Toshiba/Desktop/GGIT-1ST/3rd_semester/Thesis/Work/MapInfo/Graphs/No_Mall.3tf" + "\"Series In Columns");

Form1.mi.Do("set graph title \"No_University\"");

Form1.mi.Do("select Range_id_all_provincies, sum(No_Hospital) from thiessen_all where obj intersects o group by Range_id_all_provincies order by Range_id_all_provincies into sel4");

int p4 = tabPage3.Handle.ToInt32();

Form1.mi.Do("set next document parent " + p4.ToString() + " style 1");

Form1.mi.Do("Graph Range_id_all_provincies, COL2 From sel4 Using \"" + "C:/Users/Toshiba/Desktop/GGIT-1ST/3rd_semester/Thesis/Work/MapInfo/Graphs/No_Mall.3tf" + "\"Series In Columns");

Form1.mi.Do("set graph title \"No_Hospital\"");

Form1.mi.Do("select Range_id_all_provincies, sum(No_ATM) from thiessen_all where obj intersects o group by Range_id_all_provincies order by Range_id_all_provincies into sel5");

int p5 = tabPage5.Handle.ToInt32();

Form1.mi.Do("set next document parent " + p5.ToString() + " style 1");

Form1.mi.Do("Graph Range_id_all_provincies, COL2 From sel5 Using \"" + "C:/Users/Toshiba/Desktop/GGIT-1ST/3rd_semester/Thesis/Work/MapInfo/Graphs/No_Mall.3tf" + "\"Series In Columns");

Form1.mi.Do("set graph title \"No_ATM\"");

Form1.mi.Do("select Range_id_all_provincies, sum(No_Bank) from thiessen_all where obj intersects o group by Range_id_all_provincies order by Range_id_all_provincies into sel6");

int p6 = tabPage6.Handle.ToInt32();

Form1.mi.Do("set next document parent " + p6.ToString() + " style 1");

Form1.mi.Do("Graph Range_id_all_provincies, COL2 From sel6 Using \"" + "C:/Users/Toshiba/Desktop/GGIT-1ST/3rd_semester/Thesis/Work/MapInfo/Graphs/No_Mall.3tf" + "\"Series In Columns");

Form1.mi.Do("set graph title \"No_Bank\"");

Form1.mi.Do("select Range_id_all_provincies, sum(No_Budget_Supermarket) from thiessen_all where obj intersects o group by Range_id_all_provincies order by Range_id_all_provincies into sel7");

int p7 = tabPage7.Handle.ToInt32();

Form1.mi.Do("set next document parent " + p7.ToString() + " style 1");

Form1.mi.Do("Graph Range_id_all_provincies, COL2 From sel7 Using \"" + "C:/Users/Toshiba/Desktop/GGIT-1ST/3rd_semester/Thesis/Work/MapInfo/Graphs/No_Mall.3tf" + "\"Series In Columns");

Form1.mi.Do("set graph title \"No_Budget_Supermarket\"");

Form1.mi.Do("select Range_id_all_provincies, sum(No_Migros) from thiessen_all where obj intersects o group by Range_id_all_provincies order by Range_id_all_provincies into sel8");

int p8 = tabPage8.Handle.ToInt32();

Form1.mi.Do("set next document parent " + p8.ToString() + " style 1");

Form1.mi.Do("Graph Range_id_all_provincies, COL2 From sel8 Using \"" + "C:/Users/Toshiba/Desktop/GGIT-1ST/3rd_semester/Thesis/Work/MapInfo/Graphs/No_Mall.3tf" + "\"Series In Columns");

Form1.mi.Do("set graph title \"No_Supermarket\"");

Form1.mi.Do("select Range_id_all_provincies, sum(No_Transportation) from thiessen_all where obj intersects o group by Range_id_all_provincies order by Range_id_all_provincies into sel9");

int p9 = tabPage9.Handle.ToInt32();

```
            Form1.mi.Do("set next document parent " + p9.ToString() + " style 1");

            Form1.mi.Do("Graph Range_id_all_provincies, COL2 From sel9 Using \""
+ "C:/Users/Toshiba/Desktop/GGIT-
1ST/3rd_semester/Thesis/Work/MapInfo/Graphs/No_Mall.3tf" + "\"Series In
Columns");

            Form1.mi.Do("set graph title \"No_Transportation\"");

            this.ShowDialog();

        }

    }

}
```

**Sorted_Neighbourhoods_Houses_Pricing Form:**

```
using System;
using System.Collections.Generic;
using System.ComponentModel;
using System.Data;
using System.Drawing;
using System.Linq;
using System.Text;
using System.Threading.Tasks;
using System.Windows.Forms;

namespace Spatial_Analysis_On_Housing_Prices_Ankara
{
    public partial class Sorted_Neighbourhoods_Houses_Pricing : Form
    {
        public static MapInfo.MapInfoApplication mi;
        public Sorted_Neighbourhoods_Houses_Pricing()
        {
            InitializeComponent();
        }



        private void Sorted_Neighbourhoods_Houses_Pricing_Load(object
sender, EventArgs e)
        {
            //string ADI = string.Empty;
            //string AvG_Price_Per_Sq_meter = string.Empty;
            //string AMedian_Of_Prices_Per_Sq_meter = string.Empty;
```

```
            Form1.mi.Do("select Name , AvG_Price_Per_Sq_meter ,
Median_Of_Prices_Per_Sq_meter from Neighbourhoods order by
AvG_Price_Per_Sq_meter desc into sel noselect ");

            int range =
Convert.ToInt16(Form1.mi.Eval("int(tableinfo(sel,8)/" +
Convert.ToString("9") + ")"));

            Form1.mi.Do("fetch first from sel");

            string r1 = null; //
Convert.ToString(Form1.mi.Eval("sel.col1"));
            string r2 = null;
//Convert.ToString(Form1.mi.Eval("sel.col2"));
            string r3 =  null;
//Convert.ToString(Form1.mi.Eval("sel.col3"));
            int Cvalues = (range * 9) + 1;

            //listBox1.Items.Clear();

            for (int i = 1; i < Cvalues; i++)
            {
                Form1.mi.Do("fetch rec " + Convert.ToString(i ) + " from
sel");
                r1 = Convert.ToString(Form1.mi.Eval("sel.col1"));
                r2 = Convert.ToString(Form1.mi.Eval("sel.col2"));
                r3 = Convert.ToString(Form1.mi.Eval("sel.col3"));

                string[] row = new string[] { r1, r2, r3 };
                dataGridView1.Rows.Add(row);

                //listBox1.Items.Add(r1 + "   " + r2+"   "+r3);
            }


        }


        private void dataGridView1_SelectionChanged(object sender,
EventArgs e)
        {
            //DataGridViewRow dr  = dataGridView1.CurrentRow;
            //string ADI = dr.Cells[0].Value.ToString();
            //string AVG = dr.Cells[1].Value.ToString();
            //string Median = dr.Cells[2].Value.ToString();

            //mi = new MapInfo.MapInfoApplication();
            //int p = panel1.Handle.ToInt32();
            //mi.Do("set next document parent " + p.ToString() + "style
1");
            //mi.Do("set application window " + p.ToString());
            //mi.Do("run application \"" +
"C:/Users/Toshiba/Desktop/GGIT-
1ST/3rd_semester/Thesis/Work/MapInfo/Spatial_Analysis_on_Housing_Prices_
Ankara_City.WOR" + "\"");
```

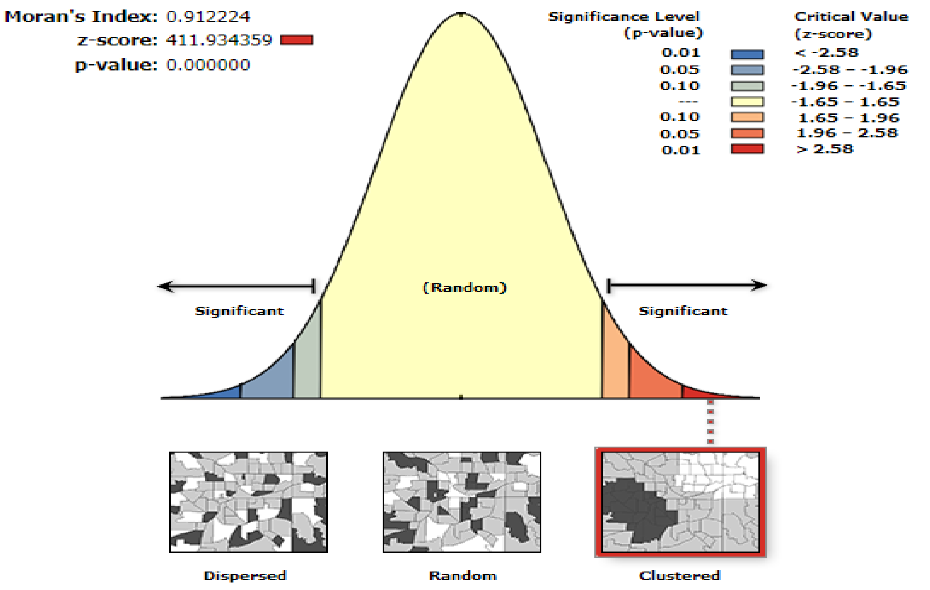
```
            //Form1.mi.Do("select * from Neighbourhoods where ADI = \""
+ ADI + "\" into sel2");
            //Form1.mi.Do("add map layer sel2 set map zoom entire layer
sel2");
        }

        private void dataGridView1_CellClick(object sender,
DataGridViewCellEventArgs e)
        {
            if (dataGridView1.SelectedRows.Count > 0)
            {
                // Create a selection query for all selected
neighborhoods
                StringBuilder selectionQuery = new StringBuilder("select
* from Neighbourhoods where ");

                for (int i = 0; i < dataGridView1.SelectedRows.Count;
i++)
                {
                    DataGridViewRow row = dataGridView1.SelectedRows[i];
                    string NAME = row.Cells[0].Value.ToString();
                    if (i > 0)
                    {
                        selectionQuery.Append(" or ");
                    }
                    selectionQuery.Append($"Name = \"{NAME}\"");
                }

                selectionQuery.Append(" into sel2");

                // Execute the selection query
                Form1.mi.Do(selectionQuery.ToString());

                // Add the sel2 layer to the map and set the zoom
                Form1.mi.Do("add map layer sel2 set map zoom entire
layer sel2");

                // Set the label font
                Form1.mi.Do("Set Map Window frontwindow() Layer sel2
Label Auto On");
                Form1.mi.Do("Set Map Window frontwindow() Layer sel2
Label Font (\"Arial\", 257, 8, 255, 16777215)");


            }
        }

    }

}
```

**CallBack Click:**

```
using System;
using System.Collections.Generic;
using System.Linq;
using System.Text;
using System.Threading.Tasks;
using System.Runtime.InteropServices;
using System.Windows.Forms;

namespace Spatial_Analysis_On_Housing_Prices_Ankara
{
    [ComVisible(true)]
    public class CallBack
    {
        Form1 f1;
        public CallBack(Form1 _f1)
        {
            f1 = _f1;
        }
        public void info(string a)
        {
            int k =
Convert.ToInt32(Form1.mi.Eval("searchpoint(frontwindow(),commandinfo(1),
commandinfo(2))"));
            string tabloadi = "";
            for (int i = 1; i <= k; i++)
            {
                tabloadi = Form1.mi.Eval("SearchInfo(" + i.ToString() +
",1)");
                String row_id = Form1.mi.Eval("SearchInfo(" +
i.ToString() + ",2)");
                Form1.mi.Do("Fetch rec " + row_id + " From " +
tabloadi);
                if ((tabloadi == "Thiessen_all"))
                {
                    f1.Invoke(new mapinfo(f1.f3.fill_form));

                }

            }


        }
        delegate void mapinfo();
    }
}
```

## D. Morans' I results Applied From 1km-5km.

### Spatial Autocorrelation Report




Given the z-score of 411.934359099, there is a less than 1% likelihood that this clustered pattern could be the result of random chance.

### Global Moran's I Summary

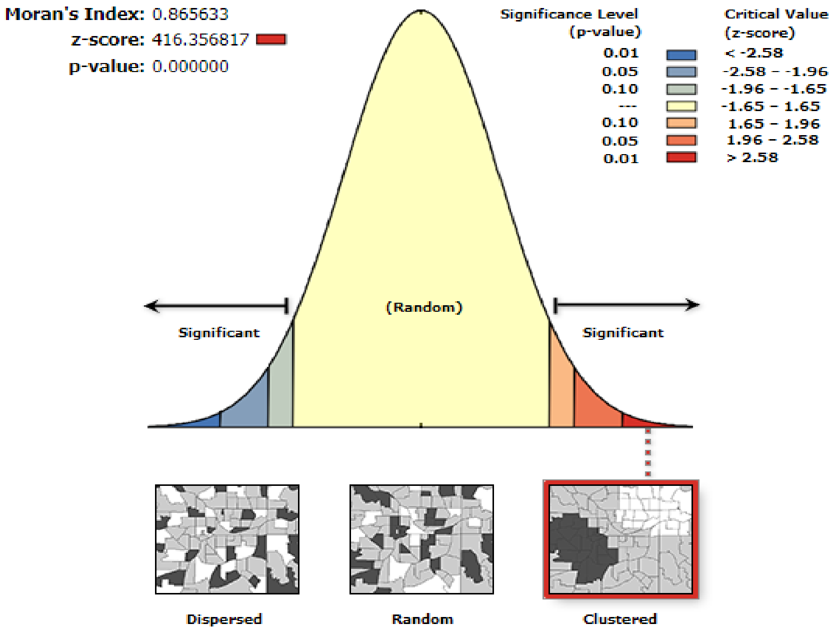
| | |
|---|---|
| **Moran's Index:** | 0.912224 |
| **Expected Index:** | -0.000024 |
| **Variance:** | 0.000005 |
| **z-score:** | 411.934359 |
| **p-value:** | 0.000000 |

### Dataset Information

| | |
|---|---|
| **Input Feature Class:** | Housing_Prices_Edited_Shifted_9Provinces |
| **Input Field:** | PRICE |
| **Conceptualization:** | INVERSE_DISTANCE |
| **Distance Method:** | EUCLIDEAN |
| **Row Standardization:** | False |
| **Distance Threshold:** | 1000.0000 Meters |
| **Weights Matrix File:** | None |
| **Selection Set:** | True |

Figure D.6: Morans'I report for 1km distance threshold.

## Spatial Autocorrelation Report




Given the z-score of 416.356817145, there is a less than 1% likelihood that this clustered pattern could be the result of random chance.

## Global Moran's I Summary

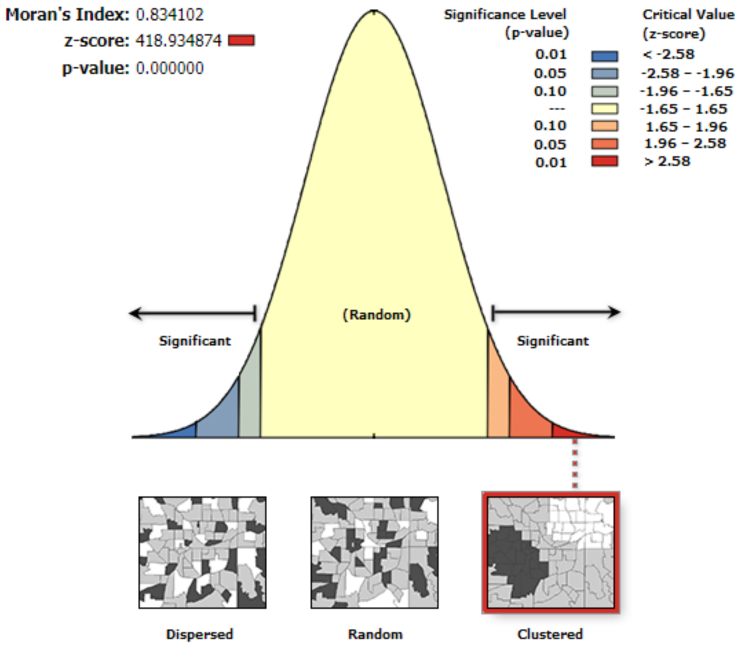
| | |
|---|---|
| **Moran's Index:** | 0.865633 |
| **Expected Index:** | -0.000024 |
| **Variance:** | 0.000004 |
| **z-score:** | 416.356817 |
| **p-value:** | 0.000000 |

## Dataset Information

| | |
|---|---|
| **Input Feature Class:** | Housing_Prices_Edited_Shifted_9Provinces |
| **Input Field:** | PRICE |
| **Conceptualization:** | INVERSE_DISTANCE |
| **Distance Method:** | EUCLIDEAN |
| **Row Standardization:** | False |
| **Distance Threshold:** | 2000.0000 Meters |
| **Weights Matrix File:** | None |
| **Selection Set:** | True |

Figure D.7: Morans'I report for 2 km distance threshold.

## Spatial Autocorrelation Report




Given the z-score of 418.934873818, there is a less than 1% likelihood that this clustered pattern could be the result of random chance.

## Global Moran's I Summary

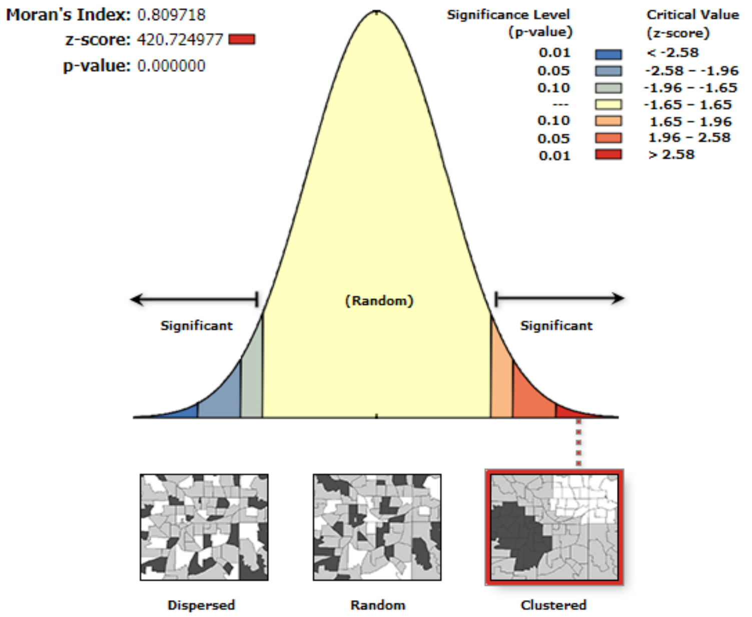
| | |
|---|---|
| **Moran's Index:** | 0.834102 |
| **Expected Index:** | -0.000024 |
| **Variance:** | 0.000004 |
| **z-score:** | 418.934874 |
| **p-value:** | 0.000000 |

## Dataset Information

| | |
|---|---|
| **Input Feature Class:** | Housing_Prices_Edited_Shifted_9Provinces |
| **Input Field:** | PRICE |
| **Conceptualization:** | INVERSE_DISTANCE |
| **Distance Method:** | EUCLIDEAN |
| **Row Standardization:** | False |
| **Distance Threshold:** | 3000.0000 Meters |
| **Weights Matrix File:** | None |
| **Selection Set:** | True |

Figure D.8: Morans'I report for 3 km distance threshold.

## Spatial Autocorrelation Report




Given the z-score of 420.724976952, there is a less than 1% likelihood that this clustered pattern could be the result of random chance.

## Global Moran's I Summary

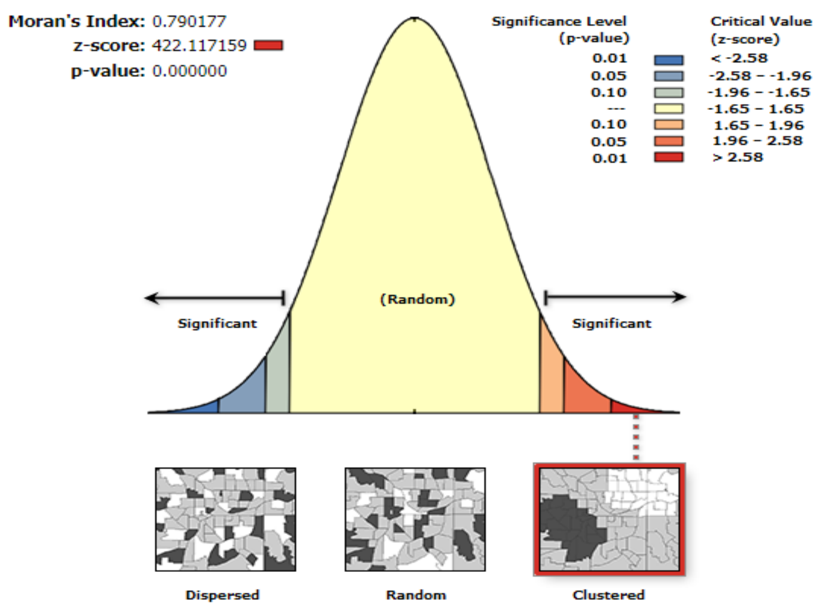
| | |
|---|---|
| **Moran's Index:** | 0.809718 |
| **Expected Index:** | -0.000024 |
| **Variance:** | 0.000004 |
| **z-score:** | 420.724977 |
| **p-value:** | 0.000000 |

## Dataset Information

| | |
|---|---|
| **Input Feature Class:** | Housing_Prices_Edited_Shifted_9Provinces |
| **Input Field:** | PRICE |
| **Conceptualization:** | INVERSE_DISTANCE |
| **Distance Method:** | EUCLIDEAN |
| **Row Standardization:** | False |
| **Distance Threshold:** | 4000.0000 Meters |
| **Weights Matrix File:** | None |
| **Selection Set:** | True |

Figure D.9: Morans'I report for 4 km distance threshold.

## Spatial Autocorrelation Report




Given the z-score of 422.117158635, there is a less than 1% likelihood that this clustered pattern could be the result of random chance.

## Global Moran's I Summary

| | |
|---|---|
| **Moran's Index:** | 0.790177 |
| **Expected Index:** | -0.000024 |
| **Variance:** | 0.000004 |
| **z-score:** | 422.117159 |
| **p-value:** | 0.000000 |

## Dataset Information

| | |
|---|---|
| **Input Feature Class:** | Housing_Prices_Edited_Shifted_9Provinces |
| **Input Field:** | PRICE |
| **Conceptualization:** | INVERSE_DISTANCE |
| **Distance Method:** | EUCLIDEAN |
| **Row Standardization:** | False |
| **Distance Threshold:** | 5000.0000 Meters |
| **Weights Matrix File:** | None |
| **Selection Set:** | True |

Figure D.10: Morans'I report for 5 km distance threshold.

## E. OLS1 and OLS2 Including Outliers Results

Table E.2: OLS1 summary results-Model variables.

| Variable | Coefficient [a] | StdError | t-Statistic | Probability [b] | Robust SE | Robust : | Robust_Pr (b) | VIP (c) |
|---|---|---|---|---|---|---|---|---|
| **Intercept** | -286252.9787 | 28548.857308 | -10.026775 | 0.000000* | 71146.263841 | -4.023444 | 0.000065* | - |
| **Area/ (m2)** | 2337.334590 | 66.734644 | 35.024306 | 0.000000* | 1203.751955 | 1.941708 | 0.052174 | 1.299277 |
| **Room Number** | 445419.48481 | 6919.596577 | 64.370730 | 0.000000* | 55216.499600 | 8.066782 | 0.000000* | 1.311189 |
| **Floor Number** | 99352.728644 | 1340.185510 | 74.133509 | 0.000000* | 3062.212333 | 32.444755 | 0.000000* | 1.118659 |
| **Near_Dist. to University** | -41.948845 | 1.738372 | -24.131113 | 0.000000* | 1.410344 | -29.743707 | 0.000000* | 1.431556 |
| **Near_Dist. to Hospital** | -73.960700 | 4.708841 | -15.706775 | 0.000000* | 4.846157 | -15.261720 | 0.000000* | 2.312142 |
| **Near_Dist. To Mall** | -19.225901 | 3.618529 | -5.313181 | 0.000000* | 3.290272 | -5,843256 | 0.000000* | 2.307250 |
| **Near_Dist. to Transport** | -505.133881 | 40.862891 | -12.361677 | 0.000000* | 61.204820 | -8.253172 | 0.000000* | 2.369706 |
| **Near_Dist. to School** | 540.794518 | 31.552093 | 17.139735 | 0.000000* | 40.026141 | 13.511033 | 0.000000* | 1.324492 |
| **Near_Dist. To ATM** | 22.111450 | 22.539631 | 0.981003 | 0.326583 | 31.048000 | 0.712170 | 0.476357 | 1.655932 |
| **Near_Dist. To Bank** | 42.870777 | 7.652492 | 5.602198 | 0.000000* | 7.464323 | 5.743425 | 0.000000* | 1.899296 |
| **Near_Dist. to Migros** | -160.025356 | 5.160426 | -31.010103 | 0.000000* | 4.882779 | -32.773418 | 0.000000* | 2.010600 |
| **Near_Dist. to Budget Supermarket** | 650.718663 | 18.542056 | 35.094202 | 0.000000* | 31.783007 | 20.473792 | 0.000000* | 2.904733 |

Table E.3: OLS2 summary results-Model variables.

| Variable | Coefficient (a) | StdError | t-Statistic | Probability [b] | Robust_SE | Robust t | Robust_Pr (b) | VIF (c) |
|---|---|---|---|---|---|---|---|---|
| **Intercept** | -397802.0861 | 28756.257908 | -13.833583 | 0.000000* | 55982.925199 | -7.105775 | 0.000000* | |
| **Room Number** | 556655.73580 | 6238.674056 | 89.226610 | 0.000000* | 14557.048386 | 38.239602 | 0.000000* | 1.034795 |
| **Floor Number** | 102938.16434 | 1355.587013 | 75.936228 | 0.000000* | 2527.746958 | 40.723287 | 0.000000* | 1.111188 |
| **Near_Dist.to University** | -43.128111 | 1.754349 | -24.583543 | 0.000000* | 1.240010 | -34.780454 | 0.000000* | 1.415535 |
| **Near_Dist. to Hospital** | -74.077450 | 4.766031 | -15.542797 | 0.000000* | 4.959340 | -14.936958 | 0.000000* | 2.299672 |
| **Near_Dist. To Mall** | -20.315607 | 3.647783 | -5.569302 | 0.000000* | 3.358191 | -6.049569 | 0.000000* | 2.276429 |
| **Near_Dist. to Transport** | -511.568559 | 41.320113 | -12.380619 | 0.000000* | 61.969442 | -8.255175 | 0.000000* | 2.352475 |
| **Near_Dist. to School** | 559.829777 | 31.702327 | 17.658949 | 0.000000* | 39.458206 | 14.187918 | 0.000000* | 1.298198 |
| **Near_Dist. To Bank** | 47.064215 | 7.622341 | 6.174509 | 0.000000* | 8.041243 | 5.852853 | 0.000000* | 1.829487 |
| **Near_Dist. to Migros** | -162.702085 | 5.208740 | -31.236362 | 0.000000* | 4.490222 | -36.234757 | 0.000000* | 1.988774 |
| **Near_Dist. to Budget Supermarket** | 672.136965 | 18.225416 | 36.879101 | 0.000000* | 31.132606 | 21.589486 | 0.000000* | 2.724652 |

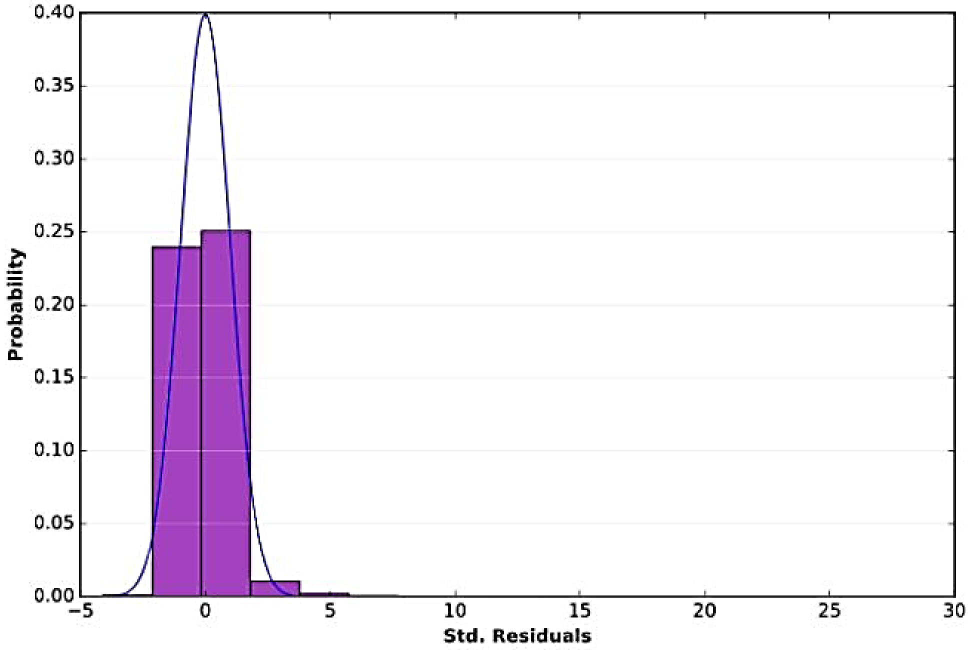



Figure E.11: OLS1 histogram of standardized residuals.

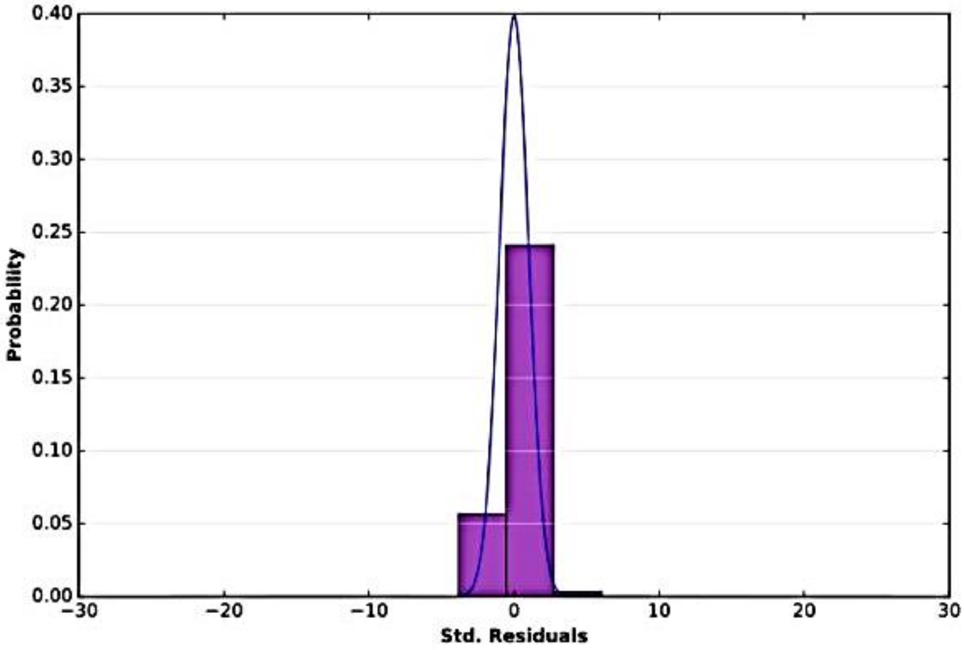



Figure E.12: OLS2 histogram of standardized residuals.

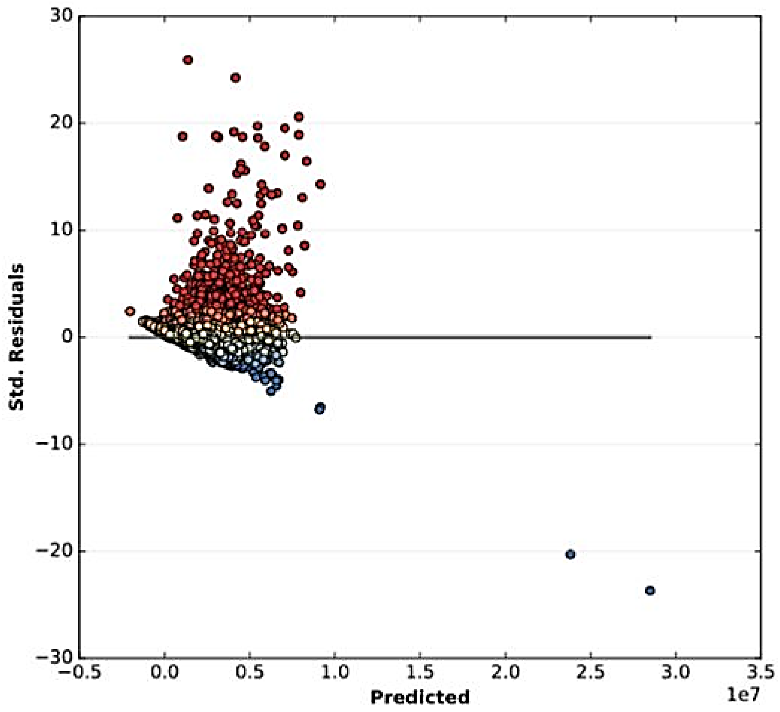



Figure E.13: OLS1 residuals vs predicted plot.

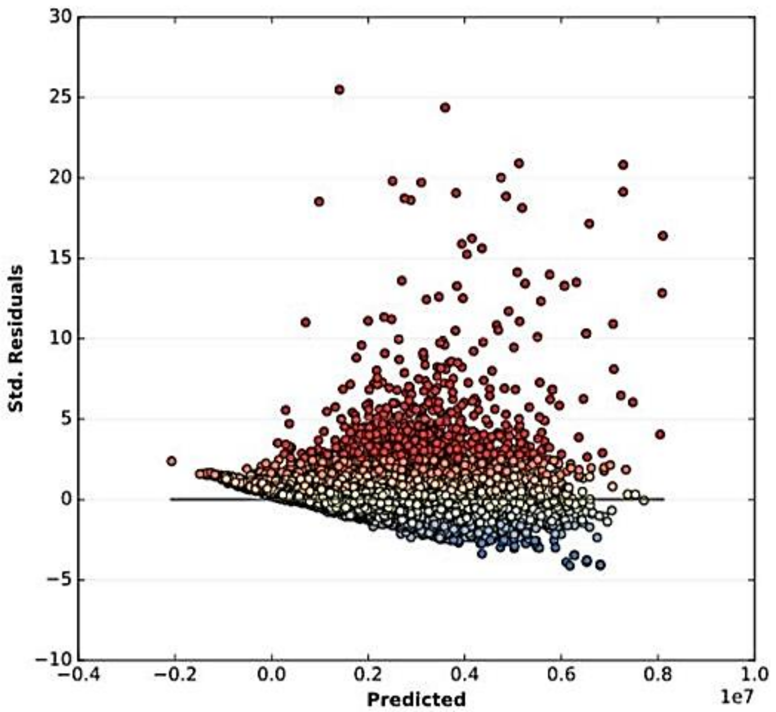
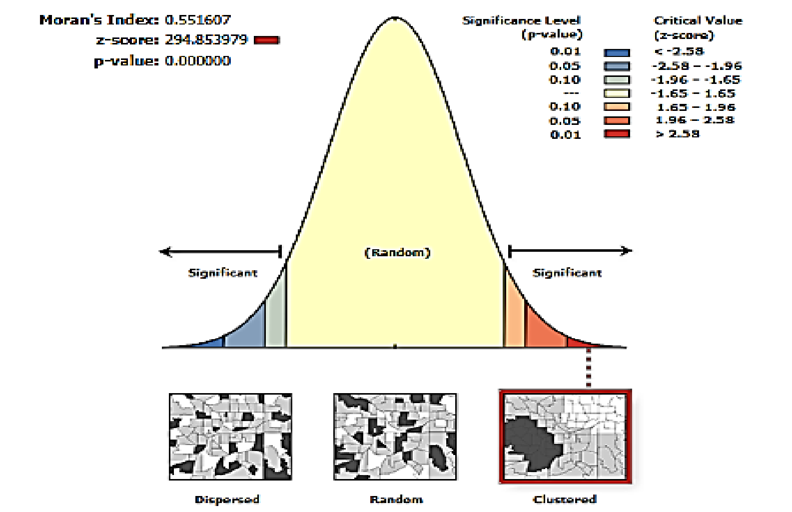



Figure E.14: OLS2 residuals vs predicted plot.

## Spatial Autocorrelation Report




Given the z-score of 294.85397939, there is a less than 1% likelihood that this clustered pattern could be the result of random chance.

## Global Moran's I Summary

| | |
|---|---|
| **Moran's Index:** | 0.551607 |
| **Expected Index:** | -0.000024 |
| **Variance:** | 0.000004 |
| **z-score:** | 294.853979 |
| **p-value:** | 0.000000 |

## Dataset Information

| | |
|---|---|
| **Input Feature Class:** | OLS_EXCLODING_AREA_AND_ATM_WITH_OUTLIERS |
| **Input Field:** | RESIDUAL |
| **Conceptualization:** | INVERSE_DISTANCE |
| **Distance Method:** | EUCLIDEAN |
| **Row Standardization:** | False |
| **Distance Threshold:** | 5000.0000 Meters |
| **Weights Matrix File:** | None |
| **Selection Set:** | False |

Figure E.15: Morans'I report for OLS2 residuals.

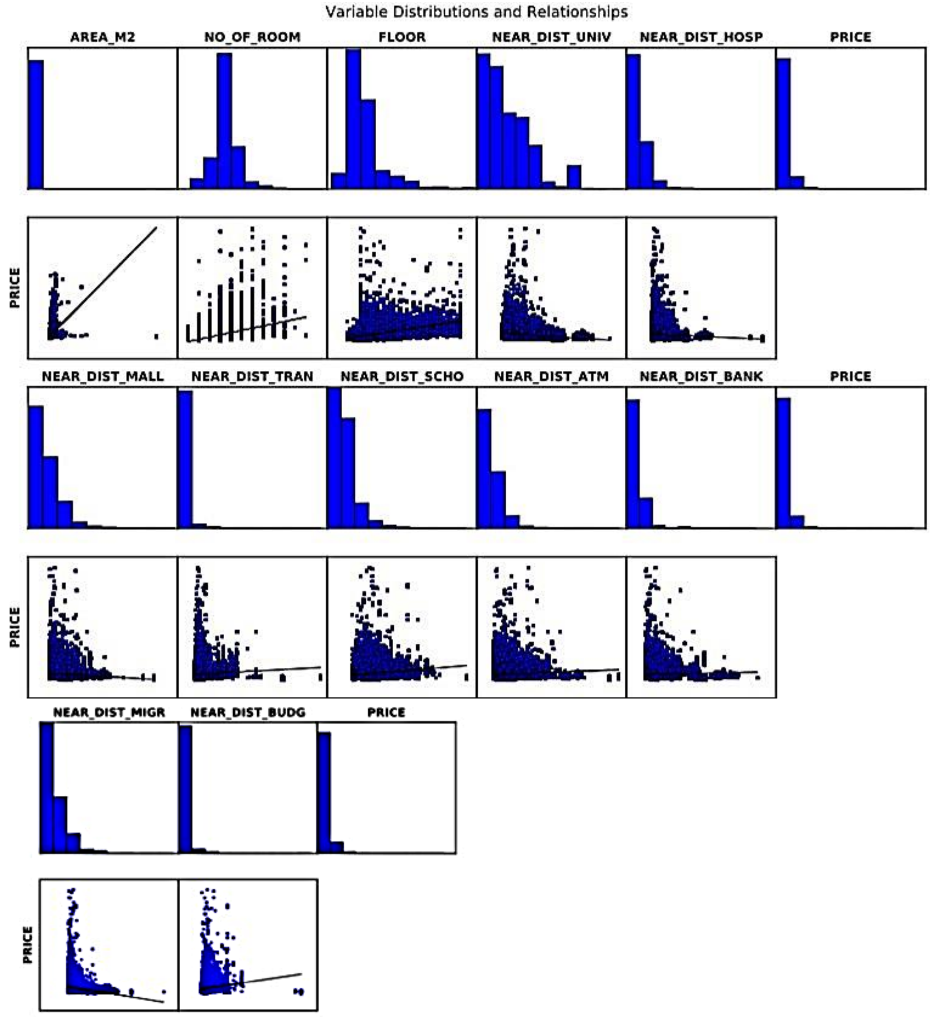



Figure E.16: OLS1 explanatory variables histogram and their relationship to the dependent variable price.

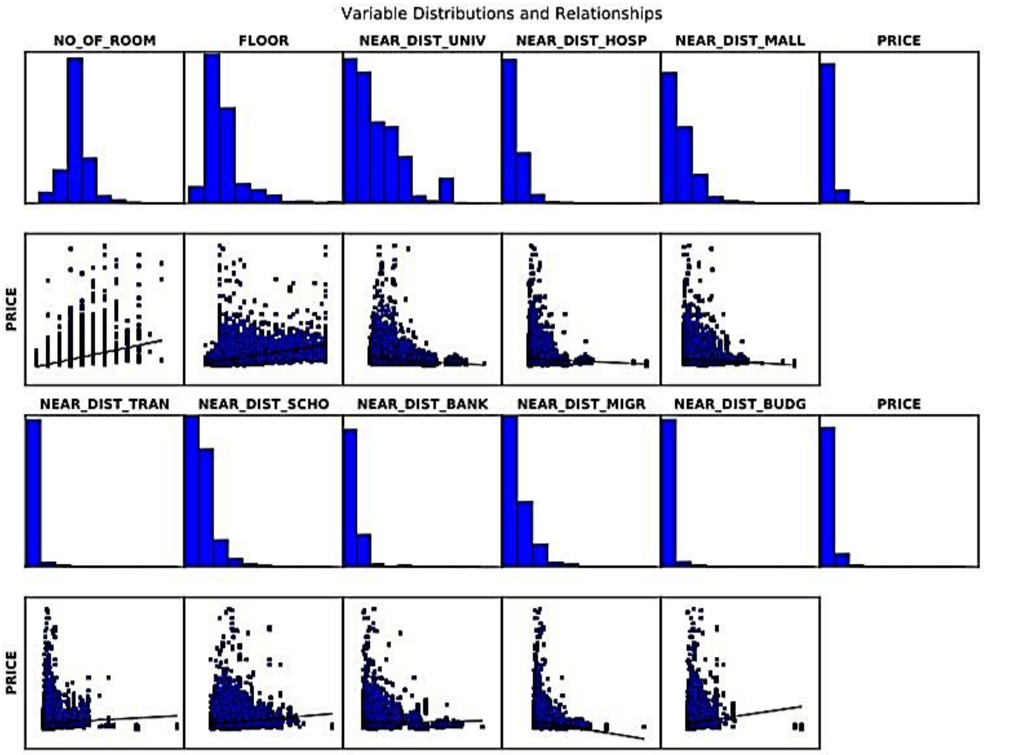



Figure E.17: OLS2 explanatory variables histogram and their relationship to the dependent variable price.

Figure E.8 illustrates the residuals' standard deviation from the OLS1 model, highlighting areas where the model's predictions deviate significantly from the observed values. This map provides a spatial representation of the model's inaccuracies, pinpointing regions where the model may be underperforming.

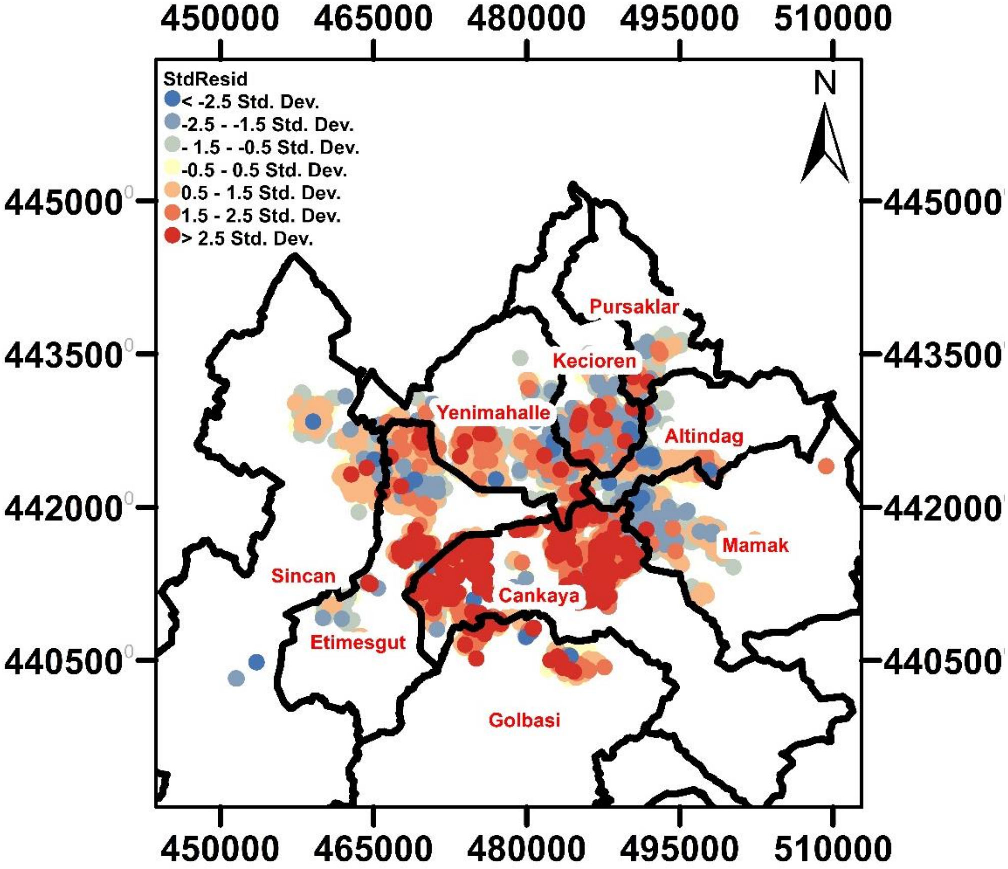



Figure E.18: OLS1 model standard residuals.

In contrast, Figure E.9 displays the standard deviation of residuals distribution for the OLS2 model. For this model, the variables deemed insufficient in the OLS1 analysis, specifically area/m² and distance to ATM, were excluded. The comparison between these maps allows for an assessment of how the exclusion of these variables impacts the model's performance across different regions. A notable improvement in the OLS2 map suggests that the removed variables were contributing to inaccuracies in the OLS1 model, thereby enhancing the overall model fit in OLS2.

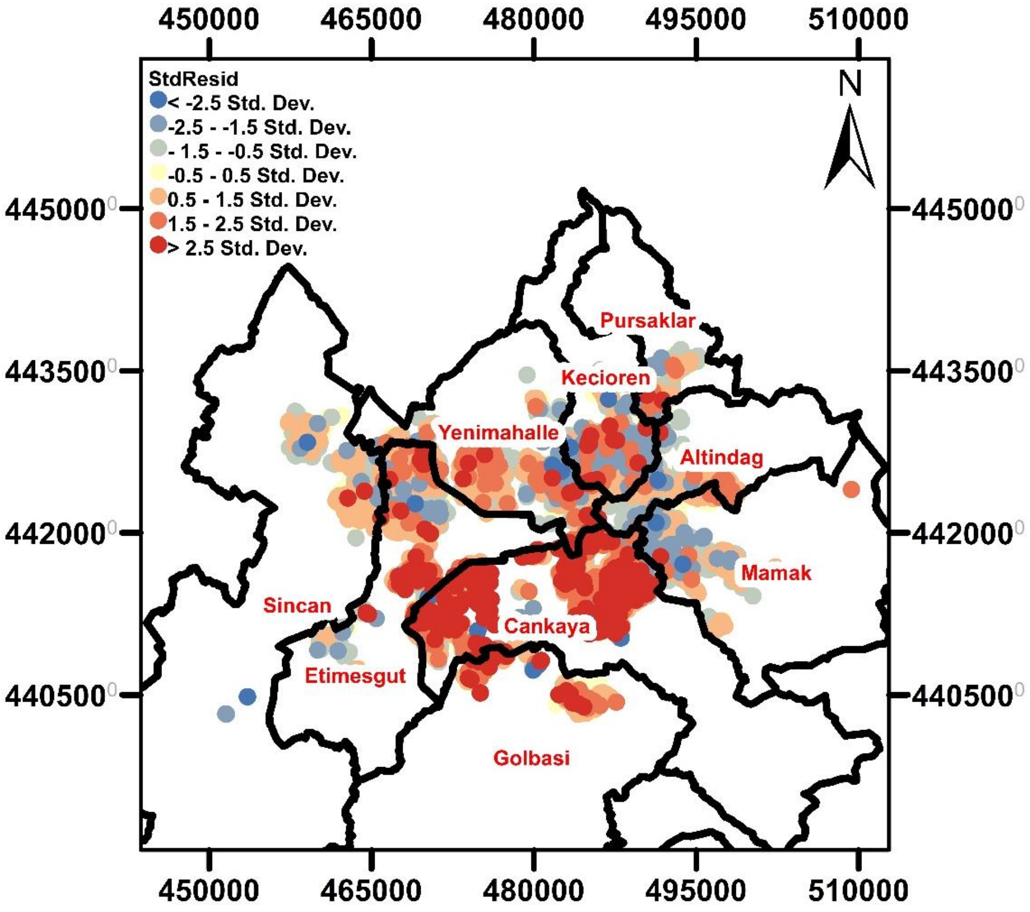



Figure E.19: OLS2 model standard residuals.

## F. GWR1 Including Outliers Results

Table F.4: The independent (Explanatory variables) used in GWR1 model (N=40910).

| Variable | Mean | Std. Dev. | Min | Q1 | Median | Q3 | Max |
|---|---|---|---|---|---|---|---|
| Price | 1825293 | 1500606.1 | 155000 | 950000 | 1399000 | 2195000 | 31900000 |
| Rooms Number | 4.1 | 0.94 | 1 | 4 | 4 | 4 | 12 |
| Floor Number | 3.4 | 4.46 | -4 | 1 | 2 | 4 | 30 |
| dist.to Transport | 177.7 | 223.5 | 8.16 | 83 | 134.3 | 204.93 | 13272.69 |
| dist.to University | 5952.7 | 7303.55 | 13.75 | 2019.91 | 4164.7 | 7450.99 | 119756.51 |
| dist.to Mall | 3765.8 | 6029.84 | 7.37 | 1533.76 | 2387.1 | 3753.2 | 119048.73 |
| dist.to Hospital | 1967.4 | 1865.47 | 4.49 | 873.19 | 1495.1 | 2551.76 | 22335.18 |
| dist.to School | 279 | 206.34 | 2.68 | 146.62 | 234.7 | 348.65 | 2297.53 |
| dist.to ATM | 374.5 | 322.97 | 0.49 | 167.5 | 305.3 | 489.62 | 3770.22 |
| dist.to Bank | 944.1 | 1018.77 | 1.93 | 403.88 | 724.4 | 1150.86 | 12031.68 |
| dist.to Migros | 1759.6 | 1554.38 | 1.15 | 611.1 | 1285 | 2480.4 | 17369.85 |
| dist.to Budget Supermarket | 293.5 | 519.97 | 0.35 | 112.97 | 187.6 | 309.61 | 8625.8 |

Table F.5: GWR1 results using 100 neighbors.

| Var. Type | VAR. | R2 | R2Adj | AICc | Neighbors |
|---|---|---|---|---|---|
| **Dependent Field** | Price | 0.91 | 0.86 | 301.85 | 100 |
| **Explanatory Field** | No_of_room | | | | |
| **Explanatory Field** | Floor | | | | |
| **Explanatory Field** | NEAR_DIST_UNIVERSITY | | | | |
| **Explanatory Field** | NEAR_DIST_HOSPITAL | | | | |
| **Explanatory Field** | NEAR_DIST_MALL | | | | |
| **Explanatory Field** | NEAR_DIST_TRANSPORTATION | | | | |
| **Explanatory Field** | NEAR_DIST_SCHOOL | | | | |
| **Explanatory Field** | NEAR_DIST_BANK | | | | |
| **Explanatory Field** | NEAR_DIST_SUPERMARKET | | | | |
| **Explanatory Field** | NEAR_DIST_BUDGET_SUPERMARKET | | | | |

Figure F.1 displays The map of the standard deviation of residuals distribution for the GWR1 model, which was applied using 100 neighbors. In this map, higher standard residuals are represented by red dots. Notably, these red dots are predominantly located in the Çankaya district. This indicates that the model is not performing as well in this area, which is known for having higher housing prices. The concentration of higher residuals suggests that the GWR1 model may be underestimating or overestimating housing prices in Çankaya, pointing to potential limitations in capturing the unique factors affecting this high-priced area.

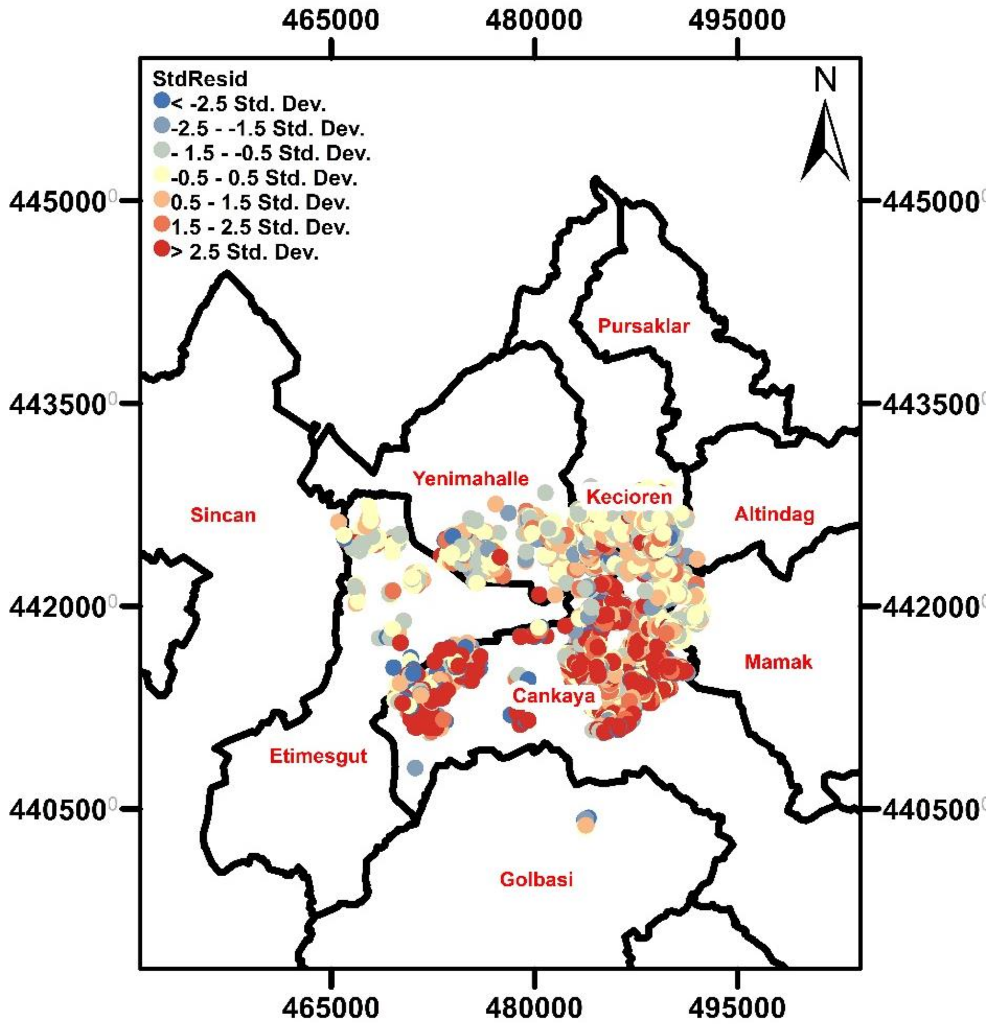



Figure F.20: GWR1 model standard residuals using 100 neighbors.

## G. Variables Histograms Excluding Outliers.

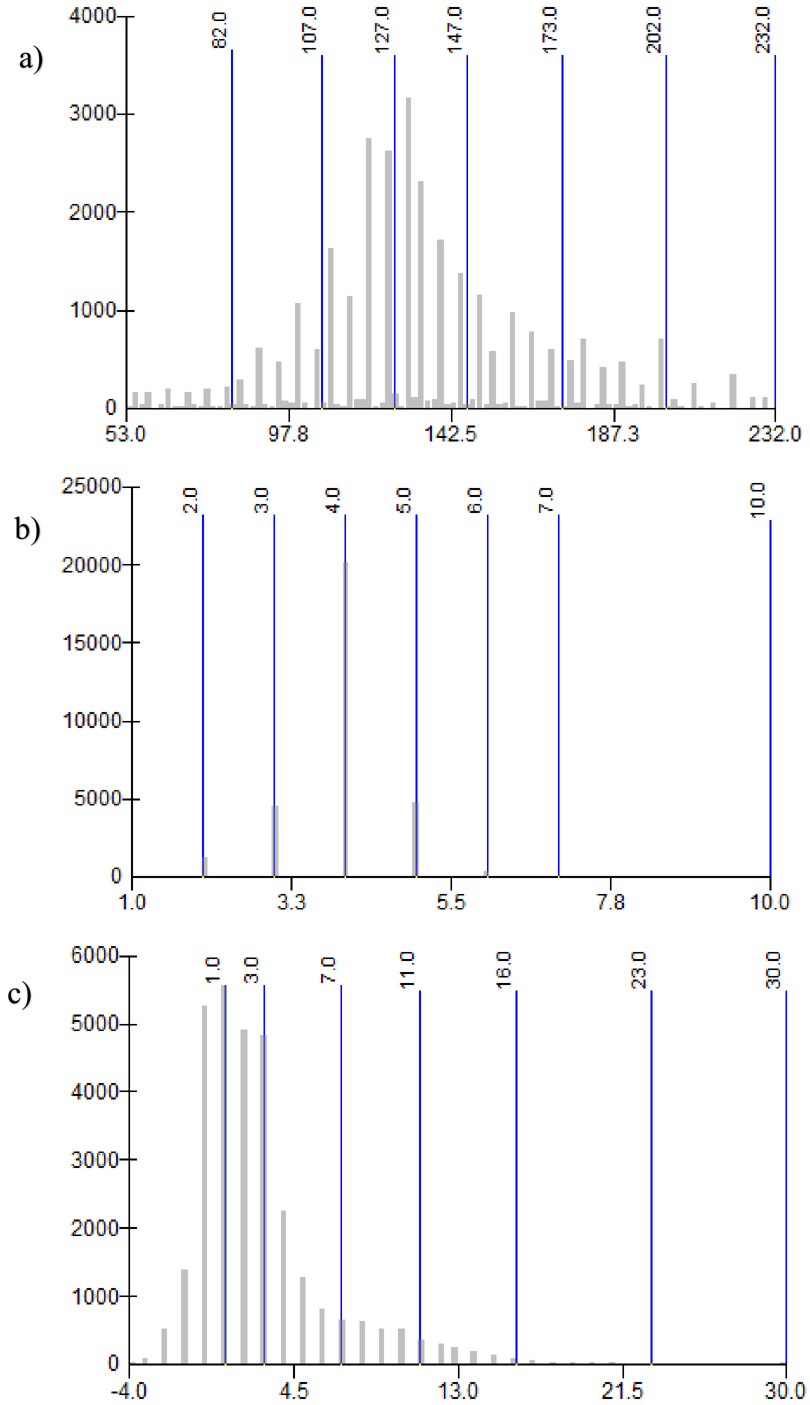



Figure G.21: Histogram excluding outliers a) Area per m2. b) Number of rooms c) Floor number.

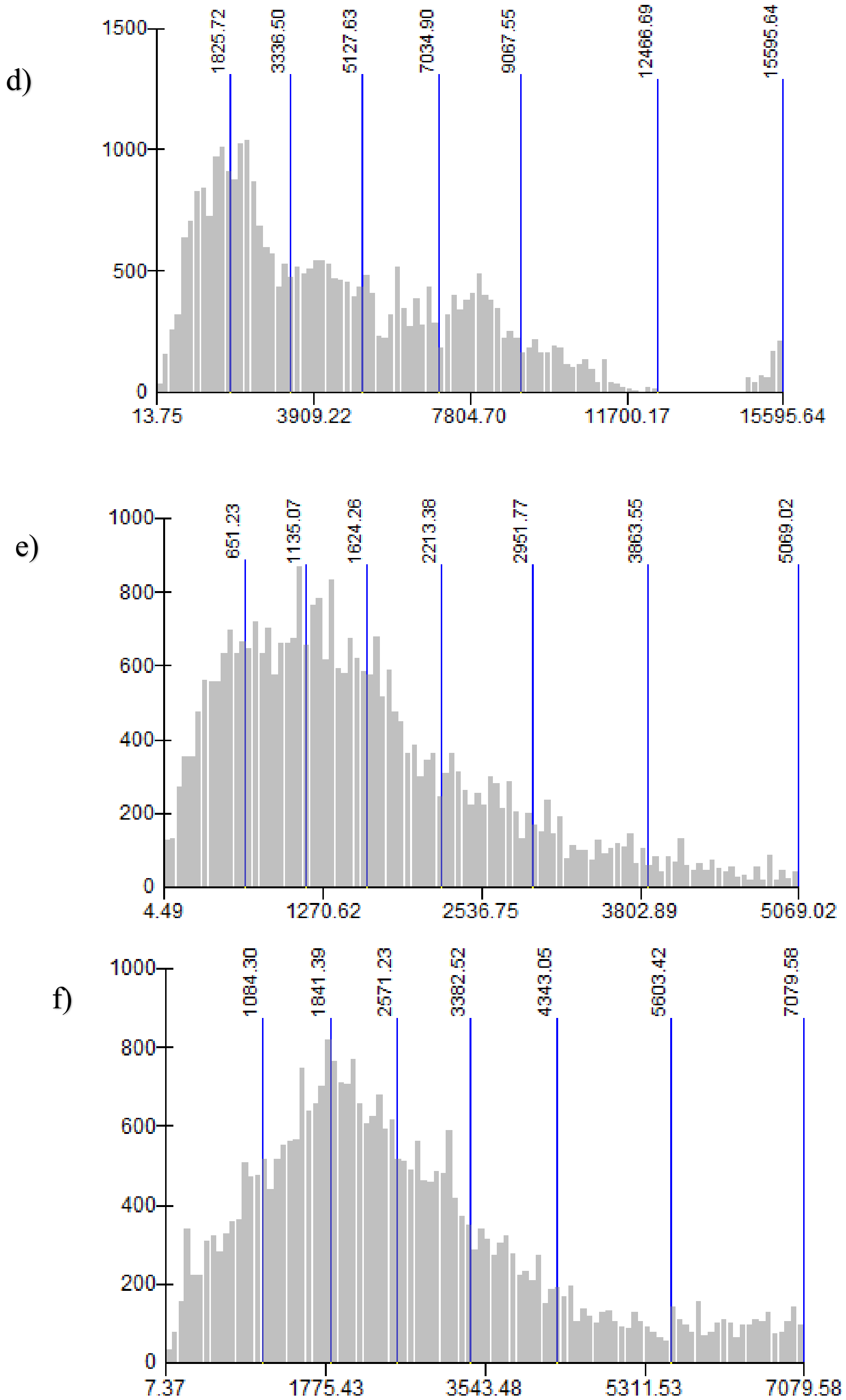



Figure G.22: Histogram excluding outliers d) Nearest distance to universities e) Nearest distance to hospitals f) Nearest distance to malls.

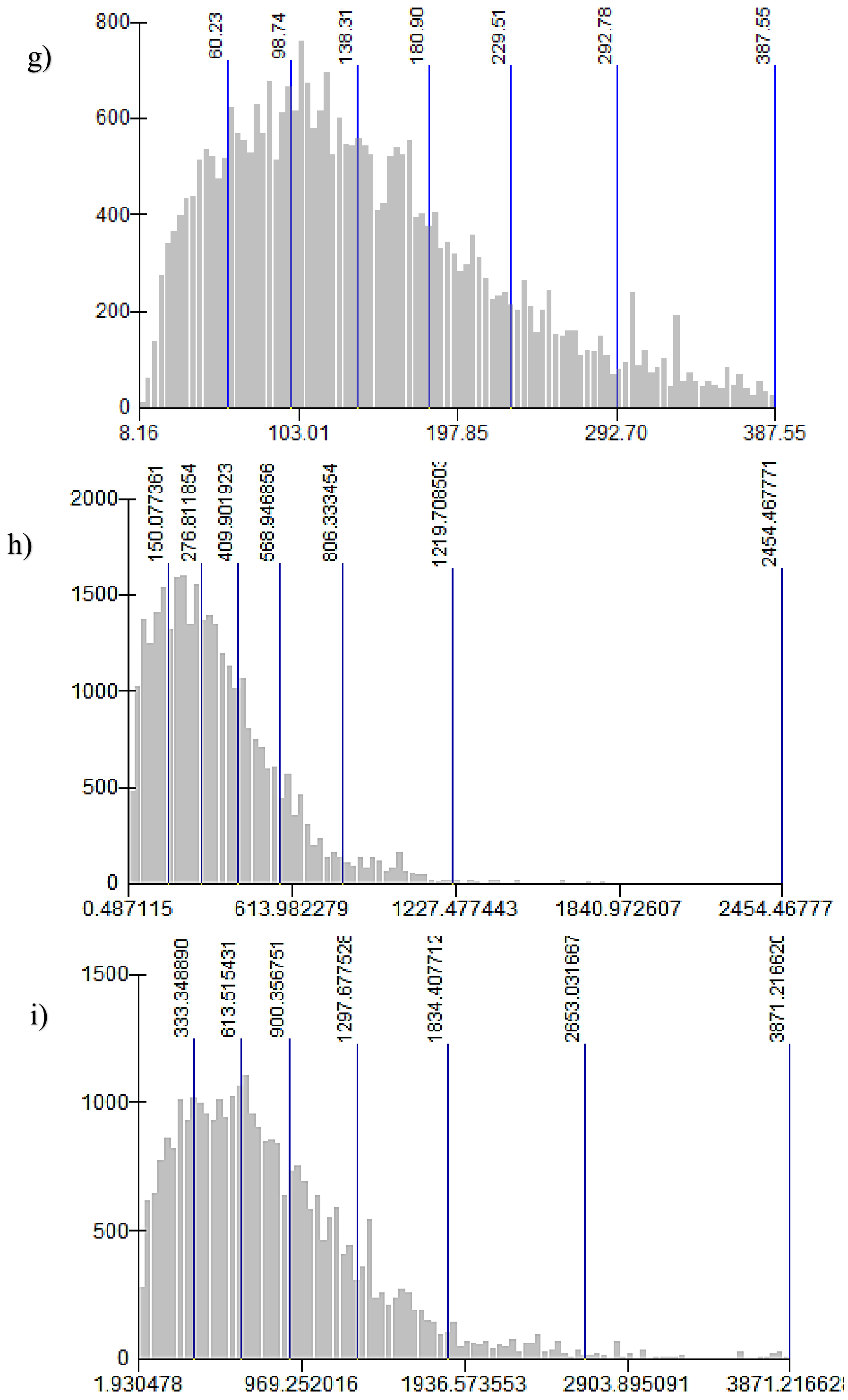



Figure G.23: Histogram excluding outliers g) Nearest distance to transportation h) Nearest distance to ATM i) Nearest distance to banks.

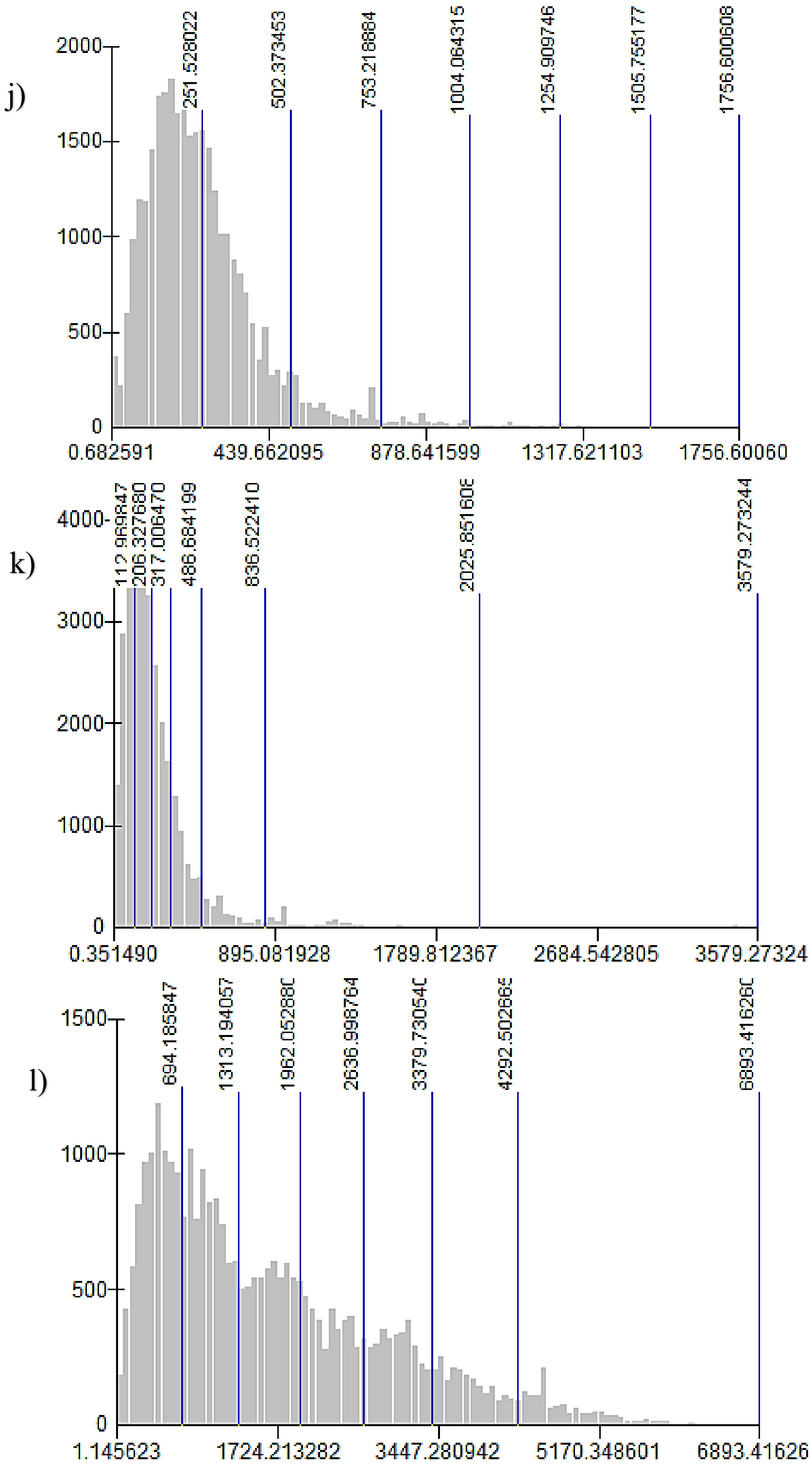



Figure G.24: Histogram excluding outliers j) Nearest distance to school k) Nearest distance to budget supermarkets l) Nearest distance to migros supermarket.

## H. OLS3 Excluding Outliers Results

Table H.6: OLS3 summery results- Model variables.

| Variable | Coefficient (a) | StdError | t-Statistic | Probability [b] | Robust SE | Robust | Robust_Pr [b] | VIF (c) |
|---|---|---|---|---|---|---|---|---|
| **Intercept** | 66427.159990 | 18668.842341 | 3.558183 | 0.000389* | 22296.967332 | 2.979202 | 0.002903* | - |
| **Area/ (m2)** | 13280.288518 | 159.713798 | 83.150540 | 0.000000* | 194.944042 | 68.123593 | 0.000000* | 3.023880 |
| **Room Number** | -93325.12210 | 6940.739976 | -13.445990 | 0.000000* | 8466.164611 | -11.023306 | 0.000000* | 2.838420 |
| **Floor Number** | 57843.428741 | 858.032464 | 67.414033 | 0.000000* | 1065.498266 | 54.287680 | 0.000000* | 1.173491 |
| **Near_Dist. to University** | -23.952312 | 1.182102 | -20.262472 | 0.000000* | 1.038665 | -23.060666 | 0.000000* | 1.604150 |
| **Near_Dist. to Hospital** | -41.349417 | 3.240966 | -12.758360 | 0.000000* | 3.143237 | -13.155044 | 0.000000* | 1.260354 |
| **Near_Dist. To Mall** | 4.118655 | 2.610841 | 1.577520 | 0.114701 | 2.715163 | 1.516909 | 0.129315 | 1.819018 |
| **Near_Dist. to Transport** | 303.371901 | 38.807580 | 7.817336 | 0.000000* | 40.134928 | 7.558800 | 0.000000* | 1.090046 |
| **Near_Dist. to School** | 315.645361 | 19.767052 | 15.968257 | 0.000000* | 21.909439 | 14.406821 | 0.000000* | 1.211308 |
| **Near_Dist. To ATM** | -7.602132 | 14.883549 | -0.510774 | 0.609525 | 16.503618 | -0.460634 | 0.645079 | 1.390636 |
| **Near_Dist. To Bank** | 83.074615 | 6.666846 | 12.460856 | 0.000000* | 7.188735 | 11.556222 | 0.000000* | 1.588717 |
| **Near_Dist. to Migros** | -154.583000 | 3.694023 | -41.846786 | 0.000000* | 3.660157 | -42.233979 | 0.000000* | 2.354437 |
| **Near_Dist. to Budget Supermarket** | 511.906577 | 16.653082 | 30.739450 | 0.000000* | 19.850951 | 25.774525 | 0.000000* | 1.356111 |

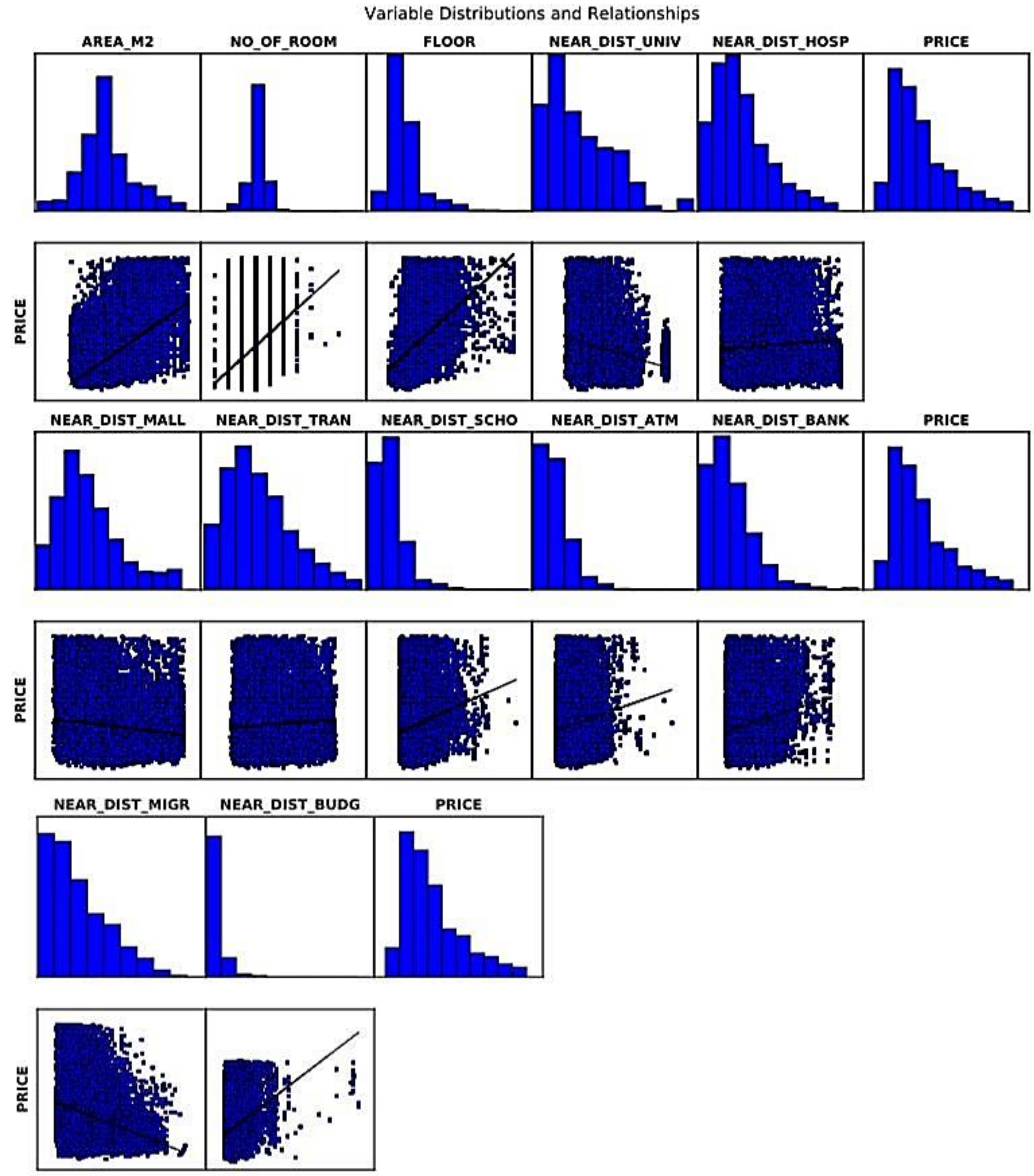



Figure H.25: OLS3 explanatory variables histogram and their relationship to the dependent variable price.

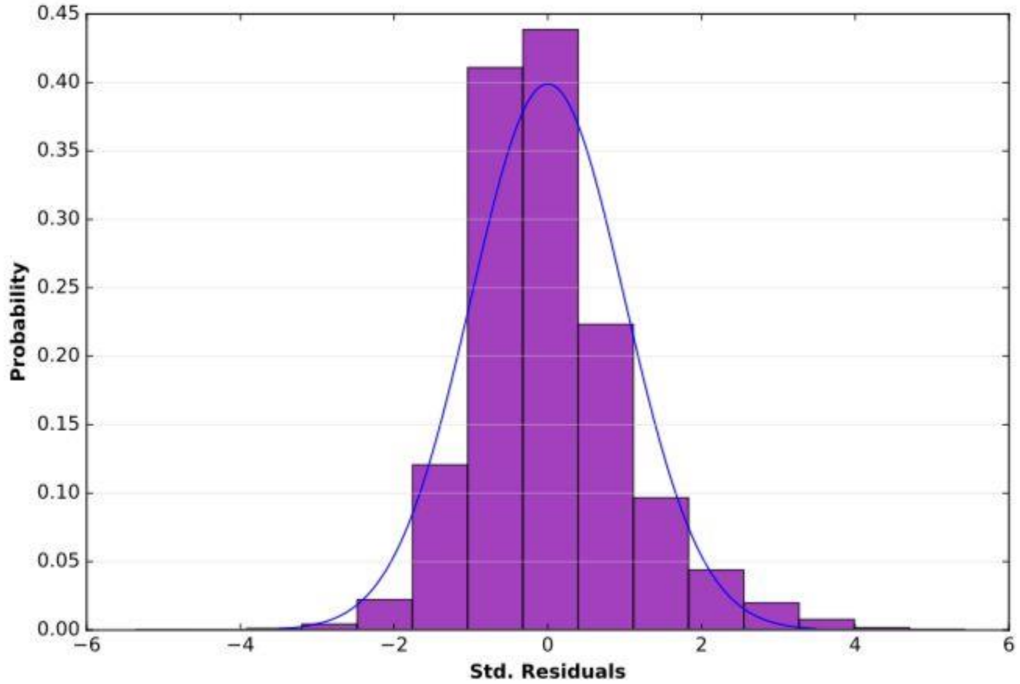



Figure H.26: OLS3 histogram of standardized residuals.

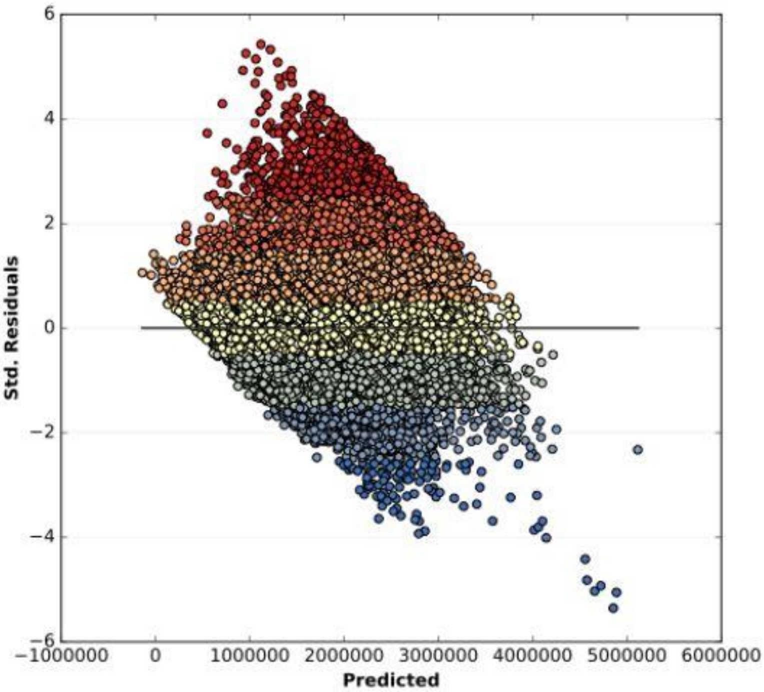



Figure H.27: OLS3 residuals vs predicted plot.

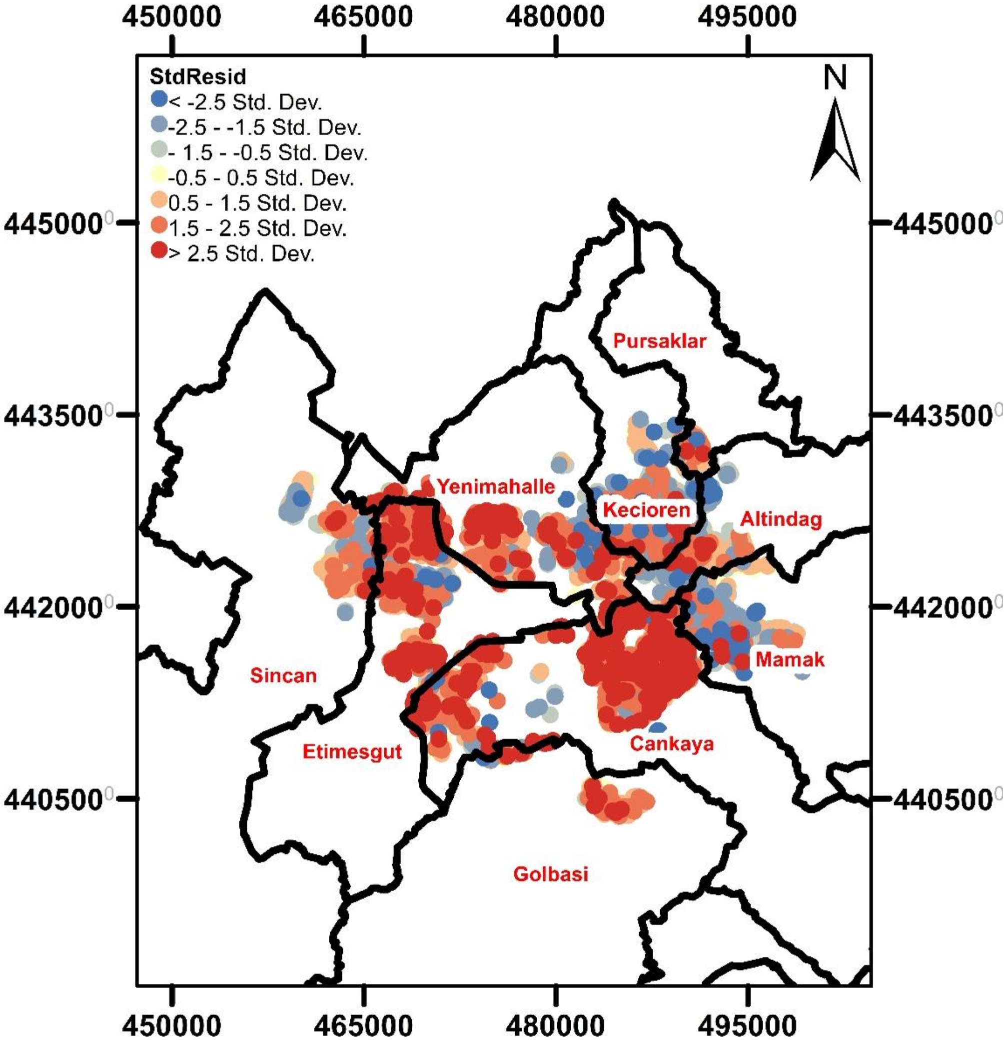



Figure H.28: OLS3 model residuals.

## I. GWR Results for Separated Dataset

### GWR3 results for Çankaya district:

Applying Geographically Weighted Regression (GWR) on the separated dataset has excluded the insufficient variables identified by Ordinary Least Squares (OLS). For the GWR3 application on the Çankaya district, the variables used are listed in Table I.1. The variables used in the GWR4 model for the other districts, excluding Çankaya, are listed in Table I.2.

Table I.7: GWR3 for Çankaya district results using 70 neighbors.

| **Var_Type** | **VAR** | **R2** | **R2Adj** | **AICc** | **Neighbors** |
|---|---|---|---|---|---|
| **Dependent Field** | Price | 0.80 | 0.71 | 1101.3 | 70 |
| **Explanatory Field** | No_of_room | | | | |
| **Explanatory Field** | Floor | | | | |
| **Explanatory Field** | NEAR_DIST_UNIVERSITY | | | | |
| **Explanatory Field** | NEAR_DIST_TRANSPORTATION | | | | |
| **Explanatory Field** | NEAR_DIST_HOSPITAL | | | | |
| **Explanatory Field** | NEAR_DIST_SCHOOL | | | | |
| **Explanatory Field** | NEAR_DIST_BANK | | | | |
| **Explanatory Field** | NEAR_DIST_MIGROS SUPERMARKET | | | | |
| **Explanatory Field** | NEAR_DIST_BUDGET SUPERMARKET | | | | |

**GWR4 results for the rest of district:**

Table I.8: GWR4 for the rest of districts results using 500 neighbors.

| Var_Type | VAR | R2 | R2Adj | AICc | Neighbors |
|---|---|---|---|---|---|
| **Dependent Field** | Price | 0.63 | 0.60 | 10748 | 500 |
| **Explanatory Field** | No_of_room | | | | |
| **Explanatory Field** | Floor | | | | |
| **Explanatory Field** | NEAR_DIST_UNIVERSITY | | | | |
| **Explanatory Field** | NEAR_DIST_TRANSPORTATION | | | | |
| **Explanatory Field** | NEAR_DIST_HOSPITAL | | | | |
| **Explanatory Field** | NEAR_DIST_SCHOOL | | | | |
| **Explanatory Field** | NEAR_DIST_BANK | | | | |
| **Explanatory Field** | NEAR_DIST_MIGROS SUPERMARKET | | | | |
| Explanatory Field | NEAR_DIST_BUDGET SUPERMARKET | | | | |

Figures I.1 and I.2 provide maps showing the distribution of the standard deviation of residuals for the GWR models applied to Çankaya and the remaining districts respectively.

It is evident from Figure I.1 and Figure I.2 that the GWR models exhibit residuals that are very closely clustered around zero. This indicates that the GWR models provide a good fit for the data in both Çankaya and the remaining districts, with residuals showing less variation and an accurate representation of spatial patterns.

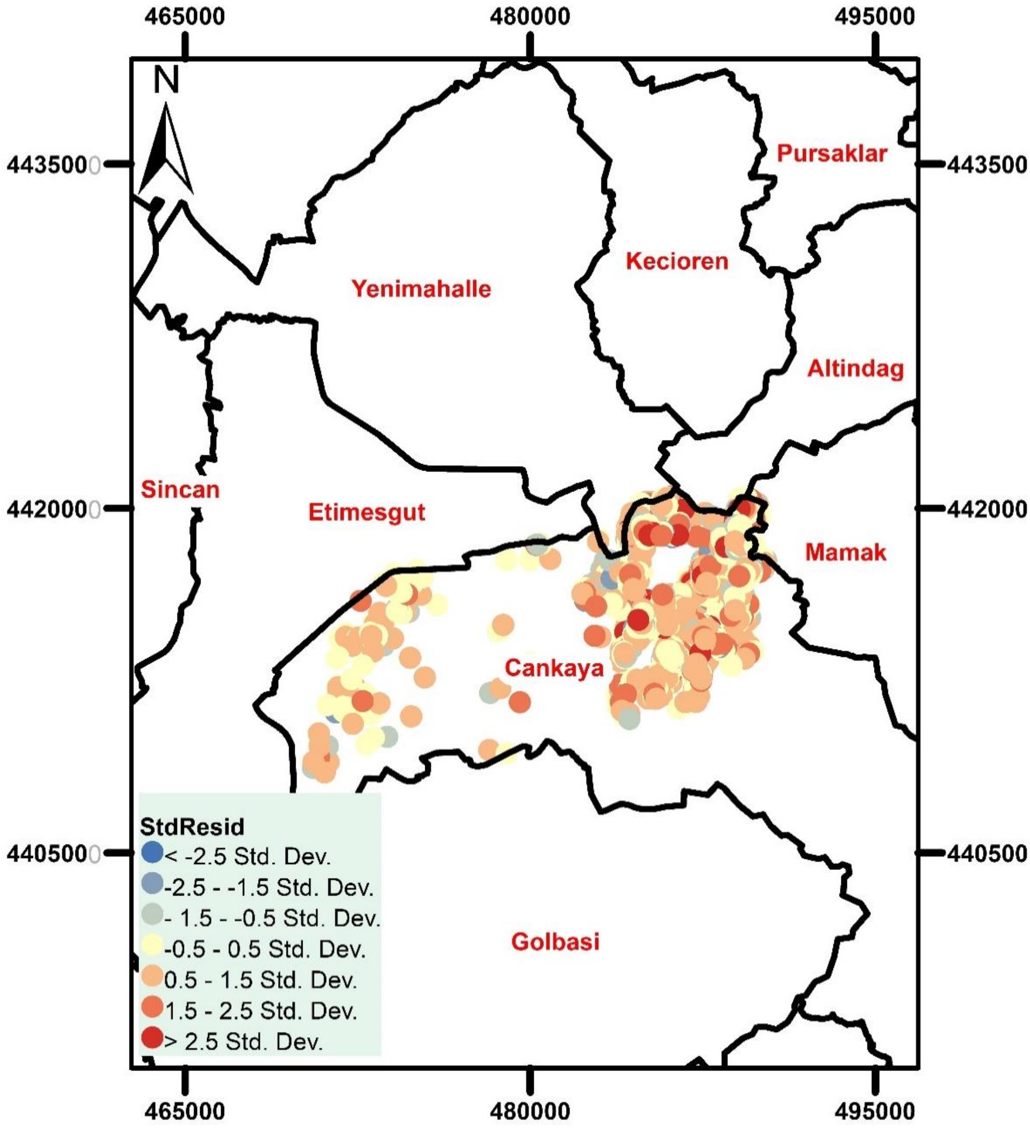



Figure I.29: GWR for Çankaya district model residuals.

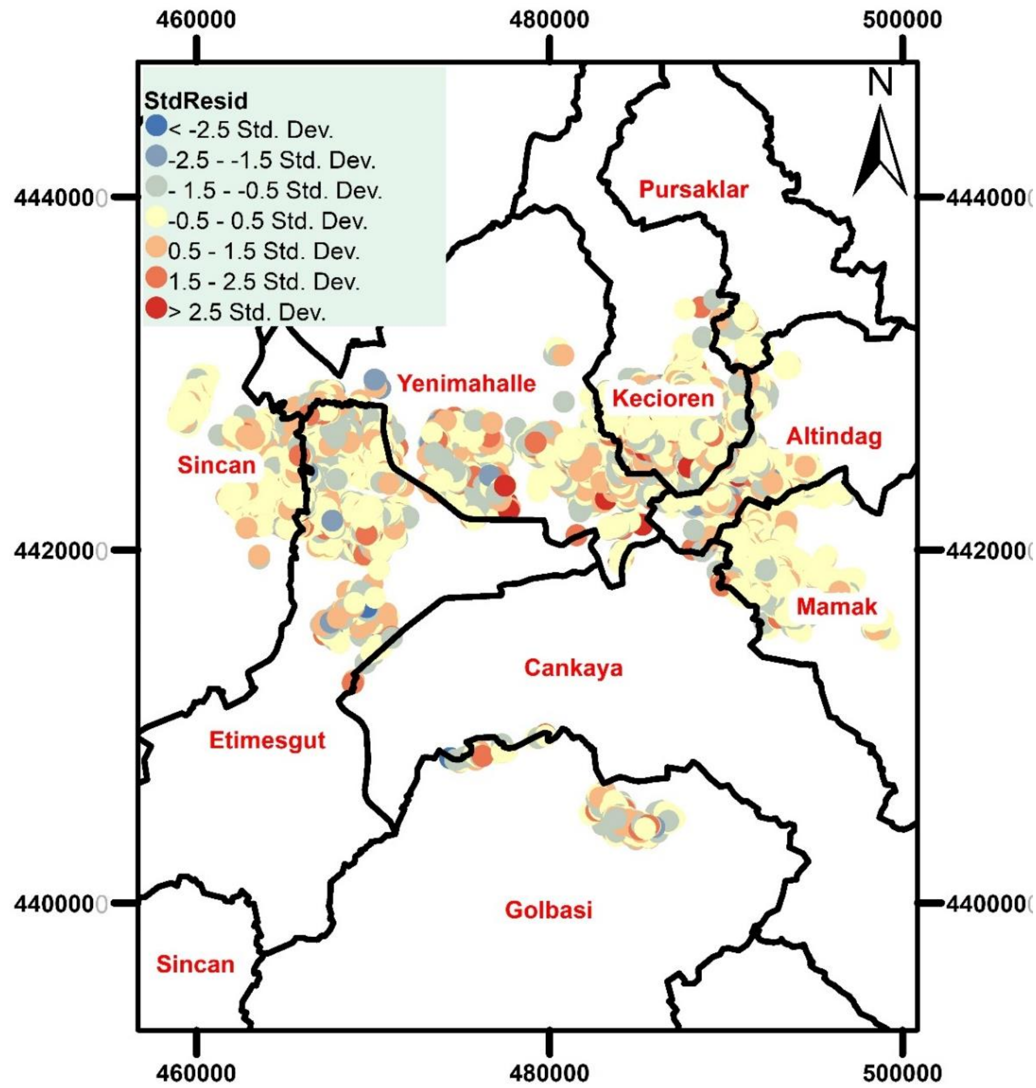



Figure I.30: GWR for rest of districts model residuals.