GIS Odyssey Journal

2022 • Volume 2 • Number 1

https://doi.org/10.57599/gisoj.2022.2.1.57

Małgorzata Renigier-Biłozor[1], Artur Janowski[2], Marek Walacik[3], Aneta Chmielewska[4]

FEATURE ENGINEERING IN PROPERTY MARKETS HOMOGENOUS AREAS DETERMINATION PROCEDURES

Abstract: Real estate is one of the most important aspect of our life and play significant role in global economy. Sooner or later, everyone has contact with properties that are place for life, work, investment, relax. That is why properties are part of many decision-making systems related to valuation, taxes, land planning and sustainable development of the areas. Analysis related to property market are based on many assumptions such as property homogeneity determination. The following paper presents proposal of utilization of automated solutions based on robust geo-estimation that enables high efficacy of property submarkets identification. The study is to propose the optimal solutions for initial part of the homogenous market analyses such as feature engineering, that enables unbiassed identification of the homogenous areas (zones). In this case the following methods based on robust geo-estimation/geoprocessing will be used: Gauss filter, geocoding and reverse geocoding, tessellation model and entropy theory.

Keywords: homogeneous areas, property market analyses, robust geo-estimation, feature engineering

Received: 16 June 2022; accepted: 21 July 2022

[1] University of Warmia and Mazury in Olsztyn, Faculty of Geoengineering, Department of Spatial Analysis and Real Estate Market, Olsztyn, Poland, ORCID ID: http://orcid.org/0000-0002-4630-7564, email: malgorzata.renigier@uwm.edu.pl

[2] University of Warmia and Mazury in Olsztyn, Faculty of Geoengineering, Department of Geodesy, Olsztyn, Poland, ORCID ID: http://orcid.org/0000-0002-5535-408X, email: artur.janowski@uwm.edu.pl

[3] University of Warmia and Mazury in Olsztyn, Faculty of Geoengineering, Department of Spatial Analysis and Real Estate Market, Olsztyn, Poland, ORCID ID: http://orcid.org/0000-0003-1904-476X, email: marek.walacik@uwm.edu.pl

[4] University of Warmia and Mazury in Olsztyn, Faculty of Geoengineering, Department of Spatial Analysis and Real Estate Market, Olsztyn, Poland, ORCID ID: http://orcid.org/0000-0001-5888-5874, email: aneta.chmielewska@uwm.edu.pl.

Introduction

Identification of comparable market areas seems to be the most neuralgic though substantial step in property analyses or valuation procedures. The one of the common interpretations of the markets states that these are, as similar to each other as possible, in the given local, conditions and types of real estates, taking into account physical-legal (endogenous) and location-economic (exogenous) aspects. The empirical problem in homogeneous markets elaboration is related to the reconciliation of two intrinsically heterogeneous phenomes: property describing factors and the properties themselves described by physical and legal characteristics. Another crucial aspect is the comparable market areas determination is the optimal way of analysis conduction.

In every property market analyses the feature selection and definition is very important, therefore their proper measurement and coding that enable their real impact reflection seems to be additional challenge in this topic.

The selection of homogeneous markets should be based on a strictly defined procedure in terms of its main stages, but as flexible as possible to ensure proper and real adaptation to the analyzed market area. An in-depth analysis of the literature on this topic was presented in the previous authors' publication entitled "Modern challenges of property market analysis-homogeneous areas" (Renigier-Biłozor et al., 2022). General stages of feature model preparation for homogenous areas determination were presented below (Fig. 1).

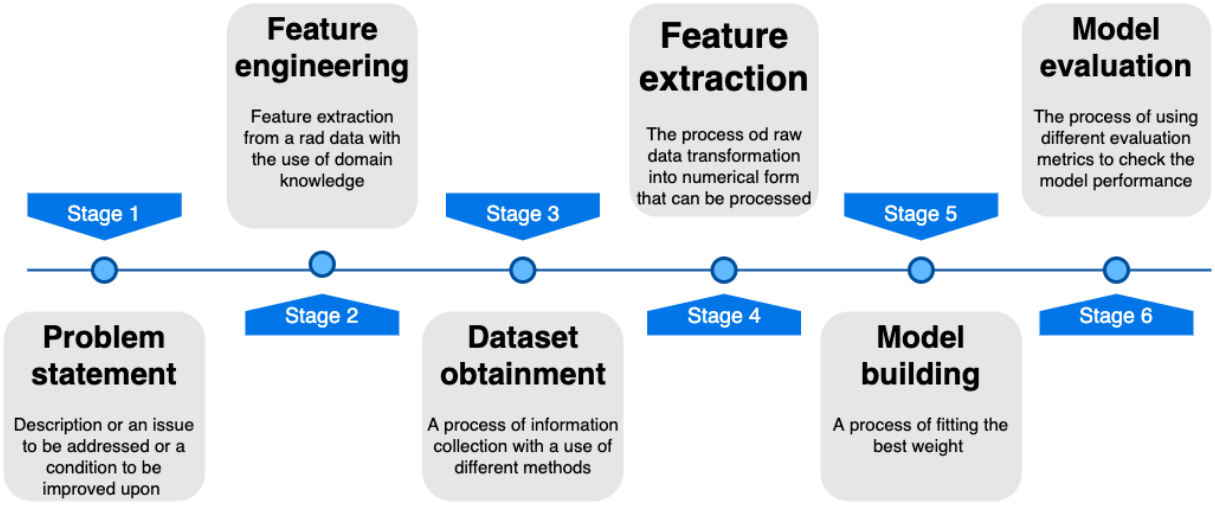

Fig. 1. Stages of feature model preparation for homogenous areas determination
Source: Own elaboration

The objective of the study is to propose the optimal solutions for initial part of the homogenous market analyses – feature engineering, that enables unbiassed identification of the homogenous areas (zones) with the use of following methods based on robust geo-estimation/geoprocessing: Gauss filter, geocoding and reverse geocoding, tessellation model and entropy theory.

Literature review and research issues

An important and indispensable stage of most market analyzes is the grouping / classification of objects into similar ones as a certain model depicting a complex / multi-dimensional phenomenon. While the details of this classification in relation to specific features or selected quantities are the domain of a given field of research, its general assumptions have (or should have) certain utilitarian schemes, stages or conditions incorporated into the technology or method used for this purpose.

When analyzing the real estate market, it should be considered that the dependencies occurring on it may be purely random, and an additional difficulty is their dynamically changing nature. Along with the growing awareness of the difficulties in getting to know the details of the processes taking place in the real estate market and reducing it "artificially" to a strictly deterministic and, in fact, perfect phenomenon, the need to look for approximate, blurred, indistinct, fuzzy solutions that can provide optimal, more satisfying results than these so-called “categorical”. The importance of the following scientific problem is underlined in the current state of art e.g. Del Giudice & De Paola (2017); Goodman & Thibodeau (1998); Manganelli et al. (2016); Morano et al. (2017); Morano & Tajani (2016); Tajani et al. (2016); Borst (2012).

In property taxation the homogeneity of the property market definition is indispensable. Usually, a taxation zone is an area in which a certain number of properties, being subject to an appraisal, demonstrate the same impact of the location on their value. According to Gnat (2019a) all the properties located in a given elementary area (a taxation zone) formally do not differ in terms of their locality. Many research work underlines the issue that indication of the zones constitutes one of the key problems that have significant impact on the accuracy of the achieved appraisal results (Cellmer & Kuryj, 2011; Manganelli et al., 2014; Morano & Manganelli, 2014; Rae, 2015; Renigier-Bilozor et al., 2019; Sawiłow, 2009; Wu and Sharma, 2012; Wu et al., 2020).

On the other hand, it is worrying that many researchers “do not understand the implication of the market segments they use in their studies and choose areas arbitrarily and ones that are too large for a meaningful evaluation of the benefit of market segmentation” and “neighborhood delineation (…) does not depend on apriori definitions such as administrative units, census tracts, ward boundaries“ (Borst, 2007). The aforementioned methods of submarkets are frequently criticized and treated as unscientific by people preferring markets classification based on empirical data (Calka, 2019; Chen et al., 2009; Mach, 2014; Salvati et al., 2019; Usman et al., 2020; Wu et al., 2018). Very interesting definition of the homogeneous areas was provided by the Eremeev et al. (2017) who claimed that objects that are not even explicitly connected and include similar objects, such as buildings, parks etc., as a rule, make up structures.

Another important issue in this topic is the determination of the size of homogeneous areas, which usually depends on the adopted assumptions. According to Manganelli at al. (2014) “the size of a homogeneous market area depends on structural factors and this value reflects the perception of market operators about location,

neighborhood, area where the property is located and inhabitants' characteristics, in practice place identity". According to Borst (2007) similarity of location is a fundamental assumption of division the universe of properties into subgroups. Numerous features of properties demand market segmentation through distinct property components consideration (Gabrielli et al., 2017; Islam & Asami, 2009; Keskin & Watkins, 2017; Warren et al., 2017).

Homogeneous areas elaboration has useful implications in terms of the property valuation, taxation, planning territorial transformations and verifying ongoing or ex post decision making. What underline Royuela and Duque after Jenkins (Jenkins, 1978; Royuela & Duque, 2013) the use of homogeneous geographic regions to define the applicability and scope of a policy or marketing strategy increase the probability of achieving the intended effects and of better predicting the unintended effects.

There is not both universal procedure and methods dedicated to homogeneous areas indication in property market analyses.

Methodology of feature engineering in homogenous areas determination

The methodology of the homogeneous property markets features engineering will be based on the several main stages. The proposed methodology is a universal and flexible solution that can be implemented in this form in other types of analyzes, aiming at homogeneous selection according to the predetermined ambiguous level of indiscernibility. In order to ensure that the analysis procedure follows the essential criteria presented in the proceeding chapters, the authors assumed the following conditions:

- the object's influence in space is not limited to a given space projection, but also includes buffers reflecting the strength of its impact,
- the division of space into optimal figures using the methodology of analytical geometry algorithms, eliminating the so-called "information dead points" at the interface of the designated area division grid (figures / units),
- "homogeneous" transactions are not strict (crisp) but a rough set with the assumed definition of similarity,
- homogeneous transactions do not have to be located only in the nearest neighborhood, the measurement of factors should be consistent with their real meaning and impact (e.g. city center travel time, sea – view, public transport accessibility etc.),
- minimizing the of the so-called behavioral simplification in the selection, quantification, and analysis of data through, inter alia, sensitivity analysis,
- classification of the attribute's significance by measuring capacity of information in data – there is no final and time-stable set of features for each type of market property.

Division of the area into optimal units.

The initial point for every spatial analysis based on vector datasets is the description of data categories that is why analytical geometry algorithms can be used for that purpose according to the following stages:

a) a rectangular node grid production at a fixed distance a.

b) creation of regular units of circles based on the node grid (that forms the circles centers).

Analysis of continuous and non-discrete impact of spatial factors for individual units according to the two simultaneous solution enables:

- provision of covering parts of neighboring circles (at assumed radius derived from analysis, with fulfilled condition that r < a), thus preliminary continuity of impact of spatial features without specified delimiting concrete skeletal lines,
- neighboring units attribute values influence increase.

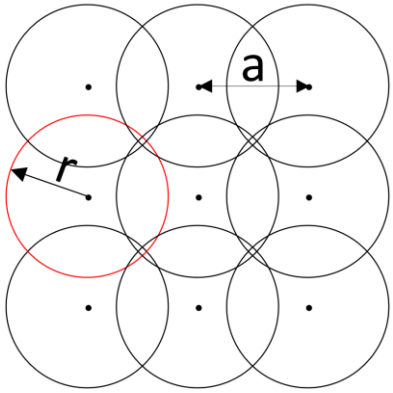

Fig. 2. Example of distances between the grid nodes and the radius of tessellation unit
Source: Own elaboration

Provided nodes grid with assigned features enable to conduct the sensitive analyses. The grid should be assumed based on the quantity of the established homogeneous groups. The nodes do not have to have strictly assigned attribute magnitude and their determination should be based on methods depending on the nature and spatial impact of the attribute. Additionally, it is worth to underline, that tessellation by means of tangent figures does not naturally consider the spatial characteristics of the neighboring units. A possible solution is to consider 4 or 8-neighborhoods, however, such a solution does not take into account the influence of units in a linearly dependent manner.

Indication and collection scope of the information about market.

In order to describe the analyzed market, the extensive literature review is usually conducted:

- **proximity and accessibility of facilities/services** is one of the factors mentioned most frequently in the literature dealing with market analysis and property valuation. The facilities include education institutions (e.g. primary schools, kindergartens etc.), health care institutions (e.g. hospitals), shopping centers (e.g.

convenience, bakeries etc.). The reason for their significance is usually connected with suitability of the neighborhood, the need and frequency of their use and time required for it. (Rabiei-Dastjerdi et al., 2021; Zheng et al., 2016);
- **environment conditions** are other characteristic indicated in the literature. Water, air or noise pollution and the accessibility of green spaces or natural hazards occurrence (e.g. floods) are the elements that seem to have the biggest influence on the residential property purchase decision. (Czembrowski & Kronenberg, 2016; Źróbek et al., 2015);
- **property communication** is another factor taken into consideration in residential property purchase decision making. The feature is usually interpreted accessibility to public transport (buts, stops, rail, metro) or distance from central business districts or public facilities (measured with time of cost of travel) (Guo et al., 2016; Hendricks, 2016; Shen & Karimi, 2017);
- **neighborhood aesthetics and social and economic background** seems to be the last predominate location characteristics described in the literature. Even though the factors might seem completely different they are usually strongly corelated. The dominating ethnicity, language, religion, family size, education level etc. usually influences the surrounding aesthetics (Bin et al., 2017; Mei et al., 2020; Pangallo et al., 2019; Xiao, 2017).

Creation of database model based on the ETL solution

Data acquisition and building a database model can be based on the classic ETL solution (Extract, Transform, Load).

Extraction

Data extraction and model determination should use of geoprocessing activities based on circle units. In order to fulfill the assumed requirements the following tasks should be conducted:
- **2D objects description** (e.g. green areas, surface area etc.), with special respect to zonal influence of individual spatial characteristics avoidance omitting – area size overlapping data layers (representing particular features);
- relevant city objects (e.g. airport or other points of interest) **distance** (e.g. euclidean) measurement for the proximity assessment of factors where the propagation or nuisance of their impact is a linear dependence on the distance or its derivative;
- **the pgRouting** use for factors proximity assessment where the propagation or nuisance of their impact is a nonlinear dependence on the distance or its derivative, distances to significant city life facilities measured along communication tracts;
- most of the data used to build the description of spatially homogeneous groups can be provided directly or indirectly form OSM (Geofabrik). OSM layers should be grouped thematically and in terms of vector representation.

Data application possibility can be analyzed, and spatial-semantic extraction performed with the use of (PostgreSQL/PostGIS/pgRouting) – Fig. 3.

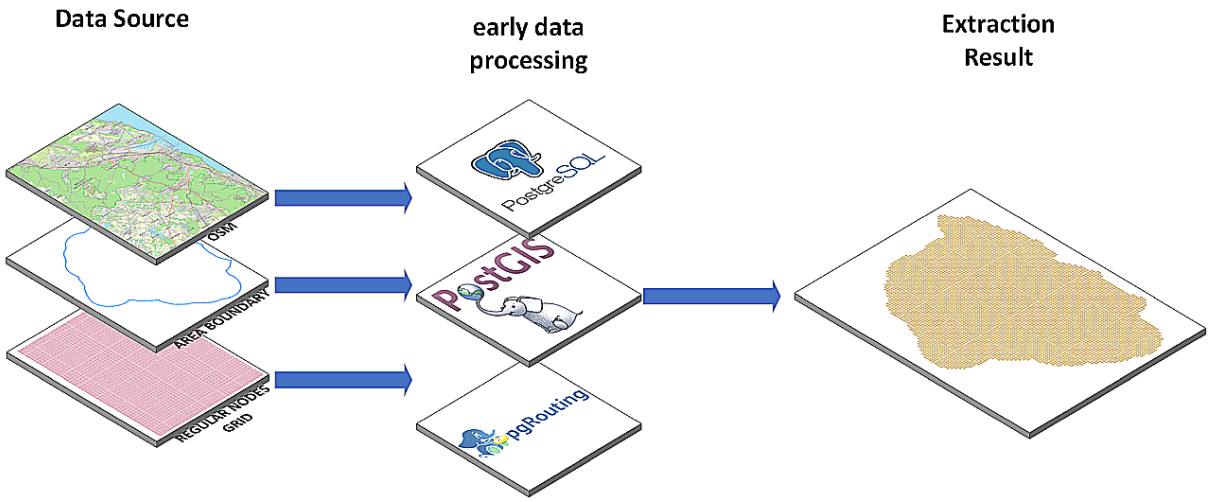

Fig. 3. Extraction process and components diagram
Source: Renigier-Biłozor et al. (2022)

As a result of the extraction and transformation of the values of all attributes, the actions taken should lead to the construction of a database whose entity-relationship diagram (ERD) example was presented below (Fig. 4).

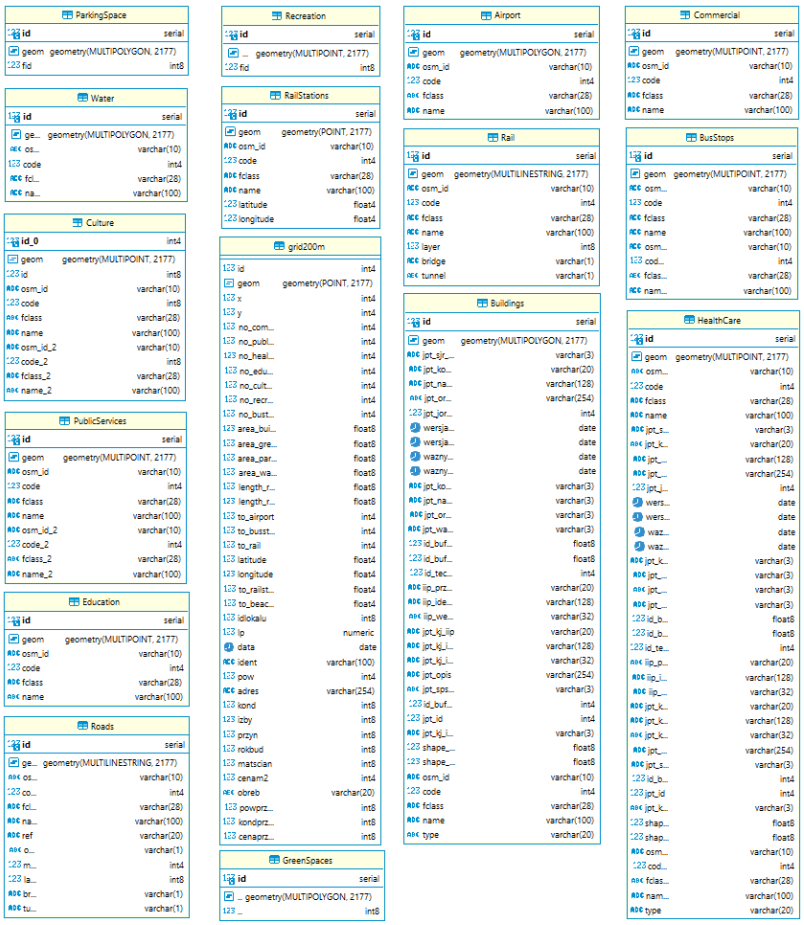

Fig. 4. Working ERD Diagram
Source: Own elaboration

Spatial extraction consisted in defining the spatial research area can be increased by a fixed buffer. Area boundary should be selected from the entire vast spatial data set only. Area boundary also makes it possible to define a tessellation layer of circular units, which in further stages of the study can have assigned quantified values of attributes describing space based on distance (Euclidean/pgRouting) to selected types of objects, their quantity or surface areas occupied by individual objects in units.

Transformation

Because of the fact that different attributes can have different characters (area, distances etc.) and type and format for storing that are not compatible – it is difficult to compare a description of numeric features, a greater unification can be indispensable. It can be processed with multiple and different types of sources data extraction and unification and smoothing. In individual units it can be obtained by the use of the entropy and Gaussian function for blurring unit tessellation.

The entropy theory allows consider simultaneously information's diversity, merit and usefulness. In the presented spin of entropy, a measure of 'disorder, chaos and randomness of certain information' (Shannon & Weaver, 1964), was used as a distinguishing factor. Entropy as a measure of homogeneity has been used in a regional analysis (Doszyń, 2008; Gnat, 2019b) as the useful measure of urban sprawl (Cabral et al., 2013), road connection order/disorder (Boeing, 2019).

Even though different theories, perceive entropy in a different way one can approach a number of equations describing it. Commonly, entropy is a logarithmic measure of the number of states with significant probability. In property market analysis the analyzed data is usually presented in the continuous form where the common formulas are inadequate. Due to this fact the following Formula 1. (Frigg and Werndl, 2011; Klein, 1970) was used for the data that consider the specificity of the geo-market information:

$$E_j = -K \sum_{i=1}^{m} n_{i,j} ln n_{i,j};\ K = 1/lnm;\ i = 1,\ldots,m;\ j = 1,\ldots,n, \qquad (1)$$

where:
Ej– entropy,
K – constant,
m – number of states in particular features,
n – expected (possible) states of feature.

Increasing the blur, stretching, mutual penetration of influence on average units can be obtained with the use of Gauss filter (especially used during blurring image data in computer vision applications and spatial data filtration (Tysiąc, 2020)) (Formula 2):

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}, \qquad (2)$$

Its discrete form of formula application for specificity of property market information is useful and efficient. Its discreet form with standard deviation σ= 1 and the size of 5×5 kernel was shown on the example below in Fig. 5.

$$G(x,y) = \frac{1}{16}$$

Fig. 5. The weights of individual neighboring units in Gauss filter discrete form for 3x3 kernel size and standard deviation equals 1
Source: Own elaboration on the OSM background

The received result data can occur to be of discrete character, which contradicts the concept of space continuity, also in terms of the real estate market. Therefore, in order to mitigate the locally high polarization of the descriptive attributes values that may occur in groups of neighboring units, Gaussian function can be used (Fig. 6).

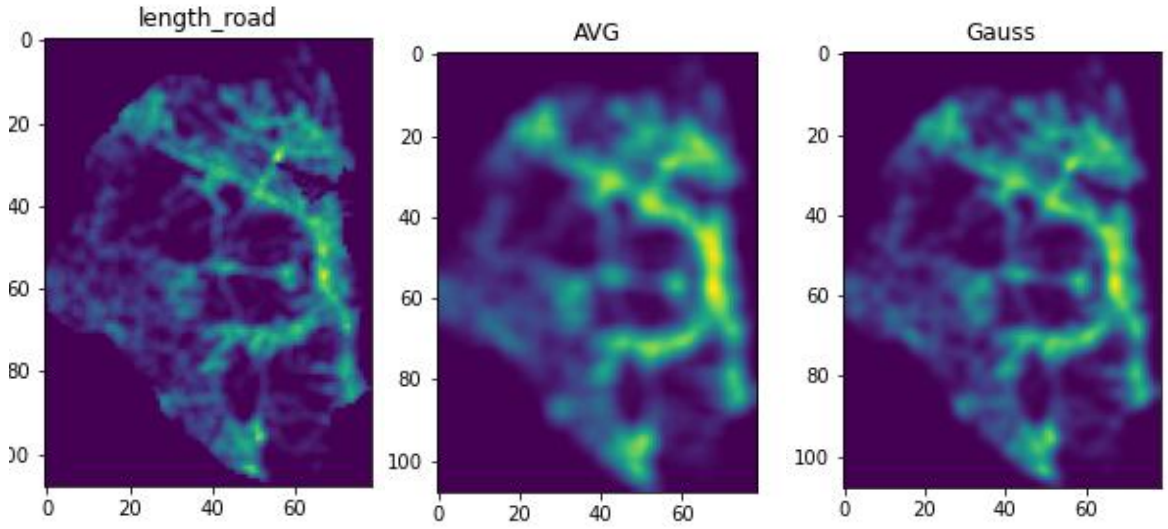

Fig. 6. Visualization of the Gauss filter attribute intensity mitigation (here: length road) with AVG calculated as arithmetic mean
Source: Own elaboration

Loading

Prepared data sets (into comparable form and assumed database) can be collected in the relational database PostgreSQL with PostGIS (spatial database extender for PostgreSQL). This gives the opportunity to:

- increase the efficiency of geometric and spatial processes with the use of spatial queries,

- automate the analyzed thematic layers geolocation of OSM, circle units and their mutual spatial relationships visualization in GIS tools (eg. QGIS, ArcGIS),
- integrate data and its portability,
- integrate data that provides compliance of the form of their storage and analysis with the objectives set up,
- multiple computing scripts (which are important in parallel way) to achieve optimal results in relation to the spatial polarization of unit's clusters.

Geo-features significance

Having removed the information redundancy an integral description of the space can be obtained within the database server – each clearly defined unit can have assigned strictly defined sets of values describing individual attributes.

For the purposes of properly features diversification in terms of their relevance and importance the entropy method can be used. This phase can be complex due to the lack of a decision factor (dependent features), to which the significance can be determined. It can be assumed, for the analysis purposes, that the relevance of variables is determined on the basis of weight of features fixed using the measures of entropy.

In this stage, a measure of entropy can be calculated for particular features according to the following Formula 3:

$$w_j = \frac{1-E_j}{\sum_{i=1}^{n}(1-E_i)}, \qquad (3)$$

where:

w_j – weight for particular feature j,

E_j – entropy calculated for particular feature j.

It must be underlined that the smaller the value of entropy and smaller variation in the analyzed information, the greater the weight of a given factor and consequently the greater the significance of the impact of the information on the outcome of the homogeneity classification/analyses. Realization of this stage enables indication of the minimum weight that should be considered in the analyses. In this stage it is possible to remove some particular data with marginal weight (under assumed significance threshold).

Conclusions

Subjectivity in property market analysis or property valuation seems to be the biggest disadvantage of the procedures and source of many misunderstandings. Therefore, while estimating property values or assessing its' investment potential, it is substantial to make the data collection and selection of homogenous groups of properties and market areas as objective and unbiased as possible. Proper comprehension of the complex real estate market dependencies requires the identification of relevant reference for analysis – submarkets. The challenge here is to determine appropriate criteria for distinguishing these submarkets and define the way of determining the similarity of the real estate market areas. The procedure of submarkets selection is multistage and involve:

- comparable area unit definition,
- particular attributes definition,
- methodology elaboration and appropriate methods selection,
- results verification.

The most troublesome, from analytical point of view, are the questions: how to combine two heterogeneous aspects - the property (physical and legal features describing it) and location (spatial features), and how to choose the size of the comparison unit?

Modern approaches to this problem are seeking for methods that allow to minimize the impact of information noise, errors or gaps in information or subjectivity in data processing and allow the processes large databases to the highest degree of raw data. Particular advantages of the presented solutions enable do diminish the following common simplifications in this area of the research:

- no need of a priori precise markets boundary definition,
- no limitations in either area or number of transaction feature description,
- no need of a priori features weights/significance definition,
- interpretation of area as contiguity phenomena described by the features elaborated in the buffer mode,
- indication of homogeneous areas as the indiscernibility areas that take into account specificity of the properties and markets related to their fundamental differentiation,
- high flexibility, scalability of the algorithm related to the boundary conditions of the model (entropy weight).

Acknowledgements

This work was supported by the National Science Centre [grant number 2019/33/B/HS4/00072, 2019/03/X/HS4/01035].

References

Bin O., Czajkowski J., Li J., Villarini G. (2017). Housing Market Fluctuations and the Implicit Price of Water Quality: Empirical Evidence from a South Florida Housing Market. Environmental and Resource Economics, vol. 68, pp. 319–341. https://doi.org/10.1007/s10640-016-0020-8.

Boeing G. (2019). Urban spatial order: street network orientation, configuration, and entropy. Applied Network Science, vol. 4, pp. 1–19. https://doi.org/10.1007/s41109-019-0189-1.

Borst R.A. (2007). Discovering and applying location influence patterns in the mass valuation of domestic real property. Doctoral thesis, Ulster: University of Ulster.

Borst R. (2012). A space-time model for computer assisted mass appraisal. Aestimum, pp. 535–545. https://doi.org/10.13128/Aestimum-13160.

Cabral P., Augusto G., Tewolde M., Araya, Y. (2013). Entropy in urban systems. Entropy, vol. 15, no. 12, pp. 5223–5236, https://doi.org/10.3390/e15125223.

Calka B. (2019). Estimating residential property values on the basis of clustering and geostatistics. Geosciences, vol. 9, pp. 1–14. https://doi.org/10.3390/geosciences9030143.

Cellmer R., Kuryj J. (2011). Określanie stref o podobnej cenności gruntów z wykorzystaniem metod geostatystycznych (*Defining the Land Value Zones with Use of Geostatistical Methods*). Studia i Materiały Towarzystwa Naukowego Nieruchomości, vol. 19, pp. 7–19.

Chen Z., Cho S.H., Poudyal N., Roberts R.K. (2009). Forecasting housing prices under different market segmentation assumptions. Urban Studies, vol. 46, pp. 167–187. https://doi.org/10.1177/0042098008098641.

Czembrowski P., Kronenberg J. (2016). Hedonic pricing and different urban green space types and sizes: Insights into the discussion on valuing ecosystem services. Landscape and Urban Planning, vol. 146, pp. 11–19. https://doi.org/10.1016/j.landurbplan.2015.10.005.

Del Giudice V., De Paola P. (2017). Spatial analysis of residential real estate rental market with geoadditive models, in: Studies in Systems, Decision and Control. Springer International Publishing, pp. 155–162. https://doi.org/10.1007/978-3-319-49746-4_8.

Doszyń M. (2008). Statystyczno-ekonometryczna analiza skłonności ludzkich (*Statistical-Econometric Analysis of Human Propensities*). Rozprawy i Studia / Uniwersytet Szczeciński, no. 707, 151 pp.

Eremeev S.V., Kuptsov K.V., Seltsova E.A. (2017). Algorithm for selecting homogeneous regions from a set of spatial objects, in: Proceedings of 2017 20th IEEE International Conference on Soft Computing and Measurements, SCM 2017. Institute of Electrical and Electronics Engineers Inc., pp. 109–112. https://doi.org/10.1109/SCM.2017.7970511.

Frigg R., Werndl C. (2011). Entropy: Aguide for the Perplexed. In: C. Beisbart, S. Hartmann (ed.), Probabilities in Physics. Oxford University Press, Oxford. https://doi.org/10.1093/acprof:oso/9780199577439.003.0005.

Gabrielli L., Giuffrida S., Trovato M.R. (2017). Gaps and Overlaps of Urban Housing Sub-market: Hard Clustering and Fuzzy Clustering Approaches. In: S. Stanghellini, P. Morano, M. Bottero, A. Oppio (ed.), Appraisal: From Theory to Practice. Green Energy and Technology. Springer, Cham. https://doi.org/10.1007/978-3-319-49676-4_15.

Gnat S. (2019a). Measurement of entropy in the assessment of homogeneity of areas valued with the Szczecin Algorithm of Real Estate Mass Appraisal. Journal of Economics & Management / University of Economics in Katowice, vol. 38, pp. 89–106. https://doi.org/10.22367/jem.2019.38.05.

Gnat S. (2019b). The Impact of the Training Set Size on the Classification of Real Estate with an Increased Fiscal Burden. Real Estate Management and Valuation, vol. 27, pp. 53–62. https://doi.org/10.2478/remav-2019-0015.

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

Guo Y., Agrawal S., Peeta S., Somenahalli S. (2016). Impacts of property accessibility and neighborhood built environment on single-unit and multiunit residential property values. Transportation Research Record: Journal of the Transportation Research Board, vol. 2568, pp. 103–112. https://doi.org/10.3141/2568-15.

Hendricks A. (2016). Christchurch, New Zealand: Proceedings. Copenhagen: International Federation of Surveyors (FIG) https://www.fig.net/resources/proceedings/fig_proceedings/index.asp (access: 05.09.2021).

Islam K.S., Asami Y. (2009). Housing market segmentation: A review of Urban and Regional Development Studies, vol. 21, pp. 93–109. https://doi.org/10.1111/j.1467-940X.2009.00161.x.

Jenkins W. (1978). Policy Analysis: A Political and Organizational Perspective. Martin Robertson, London.

Keskin B., Watkins C. (2017). Defining spatial housing submarkets: Exploring the case for expert delineated boundaries. Urban Studies, vol. 54, pp. 1446–1462. https://doi.org/10.1177/0042098015620351.

Klein M. (1970). Ludwig Boltzmann, Encyclopedia.

Mach Ł. (2014). Próba budowy homogenicznych grup województw w obszarze lokalnych rynków nieruchomości mieszkaniowych (*The attempt to create homogeneous groups of provinces in the residential real estate market*). Metody Ilościowe w Badaniach Ekonomicznych, vol. 15, pp. 219–227.

Manganelli B., De Paola P., Del Giudice V. (2016). Linear programming in a multi-criteria model for real estate appraisal, in: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Springer Verlag, pp. 182–192. https://doi.org/10.1007/978-3-319-42085-1_14.

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, vol. 9, pp. 161–177. https://doi.org/10.1504/IJBIDM.2014.065100.

Mei Y., Gao L., Zhang J., Wang J. (2020). Valuing urban air quality: a hedonic price analysis in Beijing, China. Environmental Science and Pollution Research, vol. 27, pp. 1373–1385. https://doi.org/10.1007/s11356-019-06874-5.

Morano P., Manganelli B. (2014). Estimating the market value of the building sites for homogeneous areas. Advanced Materials Research, vol. 869–870, pp. 14–19. https://doi.org/10.4028/www.scientific.net/AMR.869-870.14.

Morano P., Tajani F. (2016). Bare ownership of residential properties: Insights on two segments of the Italian market. International Journal of Housing Markets and Analysis, vol. 9, pp. 376–399. https://doi.org/10.1108/IJHMA-07-2015-0037.

Morano P., Tajani F., Locurcio M. (2017). GIS application and econometric analysis for the verification of the financial feasibility of roof-top wind turbines in the city of Bari (Italy). Renewable and Sustainable Energy Reviews, vol. 70, pp. 999–1010. https://doi.org/10.1016/j.rser.2016.12.005.

Pangallo M., Nadal J.P., Vignes A. (2019). Residential income segregation: A behavioral model of the housing market. Journal of Economic Behavior & Organization, vol. 159, pp. 15–35. https://doi.org/10.1016/j.jebo.2019.01.010.

Rabiei-Dastjerdi H., McArdle G., Matthews S.A., Keenan P. (2021). Gap analysis in decision support systems for real-estate in the era of the digital earth. International Journal of Digital Earth, vol. 14, pp. 121–138. https://doi.org/10.1080/17538947.2020.1808719.

Rae A. (2015). Online Housing Search and the Geography of Submarkets. Housing Studies, vol. 30, pp. 453–472. https://doi.org/10.1080/02673037.2014.974142.

Renigier-Bilozor M., Janowski A., Walacik M. (2019). Geoscience methods in real estate market analyses subjectivity decrease. Geosciences, vol. 9, pp. 1–20. https://doi.org/10.3390/geosciences9030130.

Renigier-Biłozor M., Janowski A., Walacik M., Chmielewska A. (2022). Modern challenges of property market analysis-homogeneous areas determination. Land Use Policy, vol. 119, pp. 106209. https://doi.org/10.1016/j.landusepol.2022.106209.

Royuela V., Duque J.C. (2013). HouSI: Heuristic for delimitation of housing submarkets and price homogeneous areas. Computers, Environment and Urban Systems, vol. 37, pp. 59–69. https://doi.org/10.1016/j.compenvurbsys.2012.04.005.

Salvati L., Ciommi M.T., Serra P., Chelli F.M. (2019). Exploring the spatial structure of housing prices under economic expansion and stagnation: The role of socio-demographic factors in metropolitan Rome, Italy. Land use policy, vol. 81, pp. 143–152. https://doi.org/10.1016/j.landusepol.2018.10.030.

Sawiłow E. (2009). Analiza wybranych metod modelowania wartości katastralnych nieruchomości (Analysis of selected methods of fixing cadastral value). Acta Scientiarum Polonorum. Geodesia et Descriptio Terrarum, vol.8, no. 2, pp. 27-38.

Shannon C.E., Weaver W. (1964). The mathematical theory of communication. The Iniversity of Illinois Press, Urbana.

Shen Y., Karimi K. (2017). The economic value of streets: mix-scale spatio-functional interaction and housing price patterns. Applied Geography, vol. 79, pp. 187–202. https://doi.org/10.1016/j.apgeog.2016.12.012.

Tajani F., Morano P., Locurcio M., Torre C.M. (2016). Data-driven techniques for mass appraisals. Applications to the residential market of the city of Bari (Italy). International Journal of Business Intelligence and Data Mining, vol. 11, pp. 109–129. https://doi.org/10.1504/IJBIDM.2016.081604.

Tysiąc P. (2020). Bringing bathymetry lidar to coastal zone assessment: A case study in the southern baltic. Remote Sensing, vol. 12, pp. 1–27. https://doi.org/10.3390/rs12223740.

Usman H., Lizam M., Adekunle M.U. (2020). Property price modelling, market segmentation and submarket classifications: A review. Real Estate Management and Valuation, vol. 28, no. 3, pp. 24-35. https://doi.org/10.1515/remav-2020-0021.

Warren C.M.J., Elliott P., Staines J. (2017). The impacts of historic districts on residential property land values in Australia. International Journal of Housing Markets and Analysis, vol. 10, pp. 66–80. https://doi.org/10.1108/IJHMA-02-2016-0015.

Wu C., Sharma R. (2012). Housing submarket classification: The role of spatial contiguity. Applied Geography, vol. 32, pp. 746–756. https://doi.org/10.1016/j.apgeog.2011.08.011.

Wu C., Ye X., Ren F., Du Q. (2018). Modified Data-Driven Framework for Housing Market Segmentation. Journal of Urban Planning and Development, vol. 144, no. 4, pp. 04018036. https://doi.org/10.1061/(asce)up.1943-5444.0000473.

Wu Y., Wei Y.D., Li H. (2020). Analyzing Spatial Heterogeneity of Housing Prices Using Large Datasets. Applied Spatial Analysis and Policy, vol. 13, pp. 223–256. https://doi.org/10.1007/s12061-019-09301-x.

Xiao Y. (2017). Hedonic Housing Price Theory Review, in: Springer Geography. Springer, pp. 11–40. https://doi.org/10.1007/978-981-10-2762-8_2.

Zheng S., Hu W., Wang R. (2016). How Much is a Good School Worth in Beijing? Identifying Price Premium with Paired Resale and Rental Data. The Journal of Real Estate Finance and Economics, vol. 53, pp. 184–199. https://doi.org/10.1007/s11146-015-9513-4.

Źróbek S., Trojanek M., Źróbek-Sokolnik A., Trojanek R. (2015). The influence of environmental factors on property buyers' choice of residential location in Poland. Journal of International Studies, vol. 7, no. 3, pp. 163–173. https://doi.org/10.14254/2071-8330.2015/8-3/13.