

Ireneusz Dąbrowski

Narodowy Bank Polski, Research and Financial Innovation Department

e-mail: ireneusz.dabrowski@nbp.pl

ORCID: 0000-0001-5353-7985

Łukasz Mach

Narodowy Bank Polski, Regional Branch in Opole

e-mail: lukasz.mach@nbp.pl

ORCID: 0000-0002-8200-4261

Łukasz Mikołajczyk

Narodowy Bank Polski, Regional Branch in Opole

e-mail: lukasz.mikolajczyk@nbp.pl

ORCID: 0000-0003-1334-7145

Arkadiusz Kuświk

Narodowy Bank Polski, Regional Branch in Opole

e-mail: arkadiusz.kuswik@nbp.pl

ORCID: 0000-0002-2741-1876

IDENTIFICATION OF HOMOGENEOUS GROUPS OF REAL ESTATE MARKETS AND THE PANDEMIC RESILIENCE OF THE CONSTRUCTION SECTOR*

DOI: 10.15611/pn.2021.4.03

JEL Classification: C38, R11, R31

© 2021 Ireneusz Dąbrowski, Łukasz Mach, Łukasz Mikołajczyk, Arkadiusz Kuświk

This work is licensed under the Creative Commons Attribution-ShareAlike 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-sa/4.0/>

Quote as: Dąbrowski, I., Mach, Ł., Mikołajczyk, Ł., and Kuświk, A. (2021). Identification of homogenous groups of real estate markets and the pandemic resilience of the construction sector. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 65(4).

Abstract: In the research process, homogeneous groups of voivodeships and the determinants of regional development were identified. The hierarchical method with Ward's algorithm and the k-means clustering method were employed for two periods, 2015 and 2018. In order to identify the determinants of development, an iterative multidimensional ranking method was used. The results of the research suggested that 16 voivodeships were divided into two groups,

* The paper presents personal opinions of the authors and does not necessarily reflect the official position of the Narodowy Bank Polski.

in which the leaders with the greatest potential A list of determinants with the highest and lowest dynamics was extracted. A statistical analysis was also performed based on the results of a survey of the economic situation in the construction industry, performed mainly to compare the period before and during the COVID-19 pandemic. The year of 2020, according to the GUS survey, was dominated by pessimism, both in the assessment of barriers to business activity and in the assessment of the current and future situation of enterprises.

Keywords: clusters analysis, real estate market, rank, potential, pandemic, COVID-19.

1. Introduction

Economic development is a goal pursued by global economies. Each region wants to develop as quickly as possible. However, when considering particular areas of the world or a particular country, it is worth noting that the pace of development in various regions may be different. The main objective of the conducted analysis was to identify the determinants of development of Polish voivodeships divided into homogeneous groups. The second stage involved the identification of the determinants influencing the growth or decline of the potential dynamics of the respective regions. This provided an opportunity to learn about qualities that are crucial for their development. The real estate market is one of the most important elements of the economy, which is characterised by certain cycles, the details of which can be found in the following studies: (Alqaralleh and Canepa, 2020; Huang, Liow, and Zhou, 2020; Łaszek and Olszewski 2018; Mach, Zmarzły, Dąbrowski, and Frącz, 2021). This may involve events described as self-perpetuating mechanisms that cause price increases that are not justified by economic and financial factors. Such a situation is usually called a speculative bubble. More on this subject, as well as the possible consequences such as a crisis, can be found in the following studies: (Mach 2019a; Wang, Wang, and Meng, 2019; Xie, Chen, and Wu, 2019). The conclusion is that the current situation in the economy is reflected in the real estate market (e.g. the level of interest rates, the level of construction, etc.), however this market has also an impact on the economic situation (a significant share in the Polish Classification of Business Activities, i.e. 6-8% according to estimates, share in local taxes, etc.) (Kucharska-Stasiak, 2016, pp. 95-102). Usually, the terms “real estate” or “property” are understood as dwellings that satisfy basic human needs. What is the real estate market? It is hard to describe this market in one sentence, mainly due to its complexity. The literature attempts to define it as a whole ecosystem of exchange and other similar relations occurring between market participants creating demand and supply (Kucharska-Stasiak, 2016, p. 40). This definition of the real estate market is a very good reference to both the definition of the real estate and the market itself. One of the criteria for its division is space. Regional markets may have their own particular characteristics, which determine their nature. The interest in the real estate market among researchers is high.

There are many studies that approach this topic in a variety of ways. An example of an analysis of the Polish real estate market can be found in (Surówka 2018), which deals with the housing deficit. In turn, the results of the study in (Łaszek, Olszewski and Waszczuk, 2017) indicate that Central and Eastern European cities are characterised by large housing deficits compared to Western European countries. The authors of the study also stress that the quality of existing resources is very poor. Both these structural features generate high demand for housing. In the paper by Fritzsche and Vandreii (2019), the authors analysed the impact of tax on real estate acquisition on the basis of transactions in the German single-family housing market. These papers address the real estate market as a larger element of the whole economy. When reviewing the available literature, one can find studies pointing to the more local or regional character of this market and review it from this angle. Another study (Łaszek, Leszczyński, and Olszewski, 2017) presents the development of the residential and commercial real estate market in Poland, also for individual voivodeships. The paper analyses the whole market as well as its regional counterparts. An overview of the real estate market grouped by voivodeships can also be found in (Malesa and Malesa, 2018). It discusses the current situation on the residential market based on average prices and the scale of demand. In other studies (Leszczyński and Olszewski, 2017), one can find the characteristics of price determinants, both on the primary and secondary (resale) market, for the 17 largest cities in Poland. The analysis of the real estate market can also be based on one local market. Another paper (Brzezicka, Łaszek, Olszewski, and Waszczuk, 2019) evaluated the residential market of the City of Warsaw, distinguishing between newly built residential buildings and the already existing assets, in order to identify the main market growth factors. Another type of study can be found in (Mach, 2014), where an attempt was made to determine homogeneous groups of voivodeships, i.e. regional markets, using cluster analysis and factor analysis. The result of the review allowed to distinguish three groups. Similar characteristics can be found in (Funke, Leiva-Leon, and Tsang, 2019). Grouping the EU Member States according to macroeconomic variables, and a comparison of the dynamics of the established groups with the dynamics calculated on the basis of data concerning the real estate market is presented in (Dąbrowski, Mach, Mikołajczyk, and Kuświk, 2020). The main conclusion of the comparison is the identification of the interrelation between the economy and the real estate market. The research also confirmed the utilitarian suitability of the calculation tools used for economic analyses. Grouping methods are applied in many areas and fields of science, for example in (Senetra and Szarek-Iwaniuk, 2019), attempting to identify and distinguish problem areas in the region of Warmia and Mazury in Poland on the basis of districts as the main taxonomic unit, and in (Carroquino, Garcia-Casarejos, and Gargallo, 2019), which analysed measures aimed to prevent climate change in the wine sector in Spain. Cluster analysis made it possible to identify four vineyard profiles.

In order to supplement the already existing research in the area of cluster analysis, economic development and the residential real estate market, this paper presents a comprehensive process of distinguishing groups of voivodeships, starting from determining the final list of diagnostic variables, through the selection of the optimal number of clusters and algorithms used, to the identification of development determinants. It should be noted that the utilitarian dimension of the identified determinants of development was the calculation of their dynamics of change, which provided an opportunity to identify those factors that have a positive and negative impact on the formation of developmental potentials.

2. Selection of diagnostic variables for the grouping process

The number of groups was determined in the research process. The period taken into account was 2015 and 2018, so the grouping was carried out twice to see how the strength of the structure achieved over three years. The time span of the data covered those years for which all the values of the diagnostic variables were available at the time of the analysis. The main aim of this article was to search for groups of voivodeships most similar to each other in terms of the socio-economic features related to the real estate market, but also describing this market. The real estate market can be described with many different variables, both those that come from the market and those that are in its vicinity, but which directly or indirectly affect its functioning. In (Foryś, 2011), there are such groups of factors determining the development of the real estate market as political, legal and organisational, demographic, and economic. Each of these factors is characterised by many variables; they can be either qualitative or quantitative. The author decided to use only quantitative variables in the research due to their more objective nature, but also because of their availability. The initial list of the diagnostic variables used in the grouping process is presented in Table 1.

Variables x_1 to x_{11} refer directly to the real estate market, while the others are related to this market and concern socio-economic characteristics. The complete list of variables contains different characteristics – some of them are expressed as average values, some are calculated as relative to others. Those that were not available in either form were calculated as per 1000 population. Variable x_{19} , a demographic coefficient per 1000 population, is not directly available in the data of the Statistics Poland Office. It was calculated as the ratio of live births per 1000 population and deaths per 1000 population. In order to select the final list of diagnostic variables, their initial statistical analysis was made. If a variable had a value below 10%, it was eliminated because its low level of variability did not sufficiently discriminate the analysed voivodeships. In terms of this criterion, the following were removed from the list: x_7 , x_8 , x_{13} , and x_{18} . The second criterion was the value of the correlation coefficient. It was decided that variables that are strongly correlated with each other (Pearson's correlation coefficient above 0.7) will be removed from the list. After analysing the correlation matrix the final list of diagnostic variables was as follows:

Table 1. List of diagnostic variables

x_1	dwellings, the construction of which has begun, for sale or rent per 1000 population
x_2	permits issued for the construction of dwellings, for sale or rent per 1000 population
x_3	completed dwellings, for sale or rent per 1000 population
x_4	number of transactions on the primary market per 1000 population
x_5	primary market – price per m ²
x_6	surface area of completed dwellings per 1000 population
x_7	average surface area of the completed dwellings
x_8	dwellings per 1000 population
x_9	new residential buildings per 1000 population
x_{10}	usable area of apartments sold on the primary market per 1000 population
x_{11}	ratio of the geodetic area of the country, built-up and urbanised residential areas to the total geodetic area (ha)
X_{12}	population density (persons/km ²)
X_{13}	number of households per 1000 population
X_{14}	average number of people in a household
X_{15}	balance of inter-provincial migration of persons of working age
X_{16}	number of registered persons of working age per 1000 population
X_{17}	number of de-registered persons of working age per 1000 population
X_{18}	marriages per 1000 population
X_{19}	demographic rate per 1000 population
X_{20}	number of entities with the Polish Classification of Business Activities, section f, division 41 per 1000 population
X_{21}	assessment of the financial situation of households – very good (%)
X_{22}	marketed production – total (thousand PLN) per 1000 population
X_{23}	unemployment rate
X_{24}	remuneration in the enterprise sector
X_{25}	average monthly disposable income per capita
X_{26}	GDP per capita

Source: own study based on the LDB database of Statistics Poland.

x_5 , x_9 , x_{11} , x_{12} , x_{19} , x_{20} , x_{21} , x_{23} , and x_{25} . All the variables, except the unemployment rate (destimulant), were considered to be stimulants. The voivodeships analysed on the basis of the selected variables showed the development of the real estate market, which applies to the analysed regions. This may translate into economic development due to the connections between the real estate market and the economy, but in this study it is understood more narrowly than economic development.

The variables indicated above allowed to create clusters being the result of the analysis carried out on the basis of their connection with the real estate market, social development, and the economy.

3. Identification of homogeneous groups of voivodeships within the adopted analysis criteria

The division of voivodeships into homogeneous groups was made on the basis of data that did not cover the pandemic period. Once the final set of variables was selected, they could be properly prepared for further analysis. The first step was standardisation, which made it possible to bring all the variables to comparability. The method most frequently used for this purpose is based on the average value and standard deviation (more on this subject can be found in (Walesiak and Gatnar, 2009)), but due to the results of comparative studies found in (Milligan and Cooper, 1988), where the authors compared eight data standardisation strategies, concluding that the best results were obtained for standardisation based on the range of variables, it was decided to standardise the variables according to the following formula (zero unitarisation):

$$x'_i = \frac{x_i - \min.x_i}{\max.x_i - \min.x_i} \quad \text{for } i = 1, \dots, n. \quad (1)$$

The next step was to choose the number of clusters to be established. The approach used was to perform the initial grouping using a hierarchical method. This allowed to create a dendrogram, which was then ‘cut off’ at the optimum point, to give the answer to the number of clusters that should be created. This kind of approach can also be found in (Mikołajczyk, 2017). The second way to determine the number of groups was by means of indexes to assess the quality of the classification. The index calculations were performed in the R software, in the clusterSim module described in (Walesiak and Dudek, 2006), and also in the source material (Walesiak and Dudek, 2019). The formulas according to which individual indices were calculated, can be found in (Walesiak, 2007). Table 2 shows the values of the calculated indexes depending on the number of clusters, which were limited for values from 2 to 10.

Table 2. Index values by number of clusters

Index	2 clusters	3 clusters	4 clusters	5 clusters	6 clusters	7 clusters	8 clusters	9 clusters	10 clusters
Caliński and Harabasz	7.510	6.282	5.986	5.475	5.358	5.348	5.435	5.277	5.262
Baker and Hubert	0.731	0.658	0.750	0.744	0.831	0.881	0.915	0.927	0.938
Hubert and Levine	0.409	0.383	0.434	0.410	0.388	0.376	0.339	0.290	0.241
Silhouette	0.309	0.204	0.196	0.148	0.123	0.140	0.125	0.102	0.082
Krzanowski and Lai	2.186	1.058	1.554	0.913	1.090	1.089	1.403	0.996	1.044
Hartigan	4.199	3.773	2.575	2.760	2.651	2.607	1.993	2.119	2.235
Gap*	0.438	0.555	0.634	0.676	0.735	0.899	1.092	1.079	1.072

* The diff values for each number of clusters higher than 0.

Source: own calculations based on the data of Statistics Poland.

On the basis of the results presented, it was concluded that the target number of clusters should be 2. This is due to the fact that the values obtained for individual indices should be interpreted as maximum values for: Caliński and Harabasz, Baker and Huber, Silhouette, Krzanowski and Lai, and as the minimum value for Hubert and Levine. For the Hartigan index, the smallest number of clusters for which the index value was less than or equal to 10, and for the Gap value, the smallest number of clusters for which the diff was greater than or equal to 0. The values in bold showed the optimal number of clusters, most indicated 2. The place of intersection of the dendrogram, a different way of determining the number of groups, was carried out on the basis of measures connected with the particular nodes. Chart 1 presents the dendrogram for the analysed voivodeships.

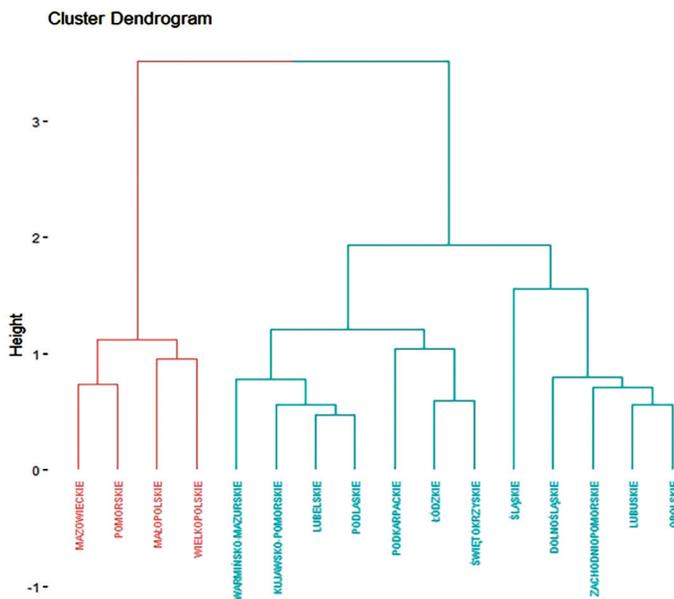


Fig. 1. Dendrogram for 2015 data with a breakdown into two clusters

Source: own study using the R software.

When determining the groups, the Euclidean distance measure and the agglomeration method with Ward's algorithm were used. This method was selected on the basis of the conclusions in (Milligan and Cooper, 1987), where the authors reviewed the methods, and where Ward's algorithm was most often mentioned as the best. The Euclidean distance was chosen because the variables taken into account were measured on a quotient scale, and could well reflect the individual distances. In order to make a comparison, a division with the k-means method was also made,

but the quality of grouping using Ward's method turned out to be better. The place of intersection of the dendrogram was determined on the basis of two measures, i.e. the difference of distance d_i and the quotient of distance q_i (Mikołajczyk, 2017). Table 3 shows the results obtained for both measures.

Table 3. Values of measures determining where the dendrogram intersects

Di	$d_i - d_{i-1}$	Qi
0.4699445	–	–
0.5577328	0.0877883	1.186805676
0.5617583	0.0040255	1.007217614
0.5995451	0.0377868	1.067265228
0.7062056	0.1066605	1.17790238
0.7369108	0.0307052	1.043479123
0.7829866	0.0460758	1.062525614
0.7989834	0.0159968	1.02043049
0.9507436	0.1517602	1.189941618
1.0380033	0.0872597	1.091780476
1.1223422	0.0843389	1.08125109
1.2034228	0.0810806	1.072242316
1.5533253	0.3499025	1.290756083
1.9300587	0.3767334	1.242533486
3.5137069	1.5836482	1.820518153

Source: own calculations.

The results, as in the case of the indexes, suggest that the dendrogram should be cut off at the height of the last level, so that two clusters are formed. This is because the maximum value of q_i must be selected. Based on the values of variables from 2015, two groups were created.

- The first group comprises the following voivodeships: Mazowieckie, Pomorskie, Małopolskie, and Wielkopolskie.
- The second group comprises the following voivodeships: Warmińsko-Mazurskie, Kujawsko-Pomorskie, Lubelskie, Podlaskie, Podkarpackie, Łódzkie, Świętokrzyskie, Śląskie, Dolnośląskie, Zachodniopomorskie, Lubuskie and Opolskie.

The basic statistics for both groups for the source values of the individual variables are presented in Table 4.

The data in Table 4 shows that group 2 is the one that has all the average values of variables lower than group 1, except for the unemployment rate, which is lower in the first group. This group is also characterised by an average negative value of the rate of natural increase per 1000 population. Taking into account the distances of individual elements of the groups from their centre, group 1 is more homogeneous.

Table 4. Basic statistical measures for both groups (2015)

Variables	Group 1	Group 2	Total
	average		
Primary market – price per m ²	5950.00	4294.42	4708.31
New residential buildings per 1000 population	2.28	1.74	1.88
Ratio of the geodetic area of the country, built-up and urbanised – residential areas to the total geodetic area (ha)	1.28%	1.04%	1.10%
Population density (persons/km ²)	2714.25	2099.75	2253.38
Demographic rate per 1000 population	1.11	0.86	0.92
Number of entities with the Polish Classification of Business Activities, section f, division 41 per 1000 population	4.75	3.43	3.76
Assessment of the financial situation of households – very good (%)	10.78%	9.00%	9.44%
Unemployment rate	7.90	11.61	10.68
Average monthly disposable income per capita	1432.58	1307.07	1338.45

Source: own calculations based on the data of Statistics Poland.

Table 5. Basic statistical measures for both groups (2018)

Variables	Group 1	Group 2	Total
	average		
Primary market – price per m ²	6068.25	4674.75	5023.13
New residential buildings per 1000 population	2.43	1.83	1.98
Ratio of the geodetic area of the country, built-up and urbanised – residential areas to the total geodetic area (ha)	1.40%	1.11%	1.18%
Population density (persons/km ²)	2616.25	2015.42	2165.63
Demographic rate per 1000 population	1.12	0.86	0.92
Number of entities with the Polish Classification of Business Activities, section f, division 41 per 1000 population	22.18%	19.67%	20.29%
Assessment of the financial situation of households – very good (%)	5.05	3.57	3.94
Unemployment rate	4.43	7.25	6.54
Average monthly disposable income per capita	1756.74	1608.09	1645.26

Source: own calculations based on the data of Statistics Poland.

Applying the same approach to the variables taken into account their standardisation, selected distance measure, and comparison of two methods, i.e. agglomeration and k-means, the process of grouping for individual voivodeships was repeated. In this case, data from 2018 were used. The indexes also suggested two clusters, while grouping was carried out on the basis of the agglomeration method (Ward's), which, as in the case of the data from 2015, was characterised by slightly better quality compared to the k-means method. The division of voivodeships turned out to be exactly the same as in 2015. Considering the average distance of each element of a given group from its centre, group 1 had a lower average and was therefore more concentrated. Analysing the two years studied, it can be concluded that the quality of grouping was similar for both periods. In 2018, the average distance of individual voivodeships from its centre in group 1 was slightly lower than in 2015, while in group 2 an opposite dependence was observed. Finally, it can be concluded that the structure of the groups was constant and stable over time and did not change over the 3-year period. Changes in the regions were expressed through the analysed characteristics related to the real estate market, which is actually a market where changes require more time.

The basic statistics for both groups for the source values of the individual variables from 2018 are presented in Table 5.

4. Identification and analysis of the dynamics of changes in socio-economic development potentials

In the further part of the work the author decided to use the created groups and their elements to estimate the potential based on the development tendency. Diagnostic variables were used in further analysis to determine the value of the aggregate index for both analysed periods as well as for 2016 and 2017 in order to determine the average rate of change of this index. This made it possible to rank the voivodeships by their value and thus to determine the potential of the analysed regions, as the maximum value of the index could be 100 for each of them. The ranking was carried out based on a linear ordering method. All the diagnostic variables were divided into stimulants and destimulants, and then for each of the voivodeships, the aggregate index was calculated. Each voivodeship was ranked on the basis of the obtained results. More on the method itself can be found in (Sokołowski and Markowska, 2017). The ranking of the voivodeships for 2015, 2016, 2017, and 2018 and the values of the calculated index are shown in Table 6.

In four voivodeships (two from group 1 and two from group 2), positive dynamics could be observed, while in the remaining ones the dynamics was negative. The highest average increase of the index in group 1 was recorded in the Pomorskie voivodeship, while in group 2 the highest increase was in the Świętokrzyskie voivodeship. It was also noted that the highest growth did not apply to the group leaders, so if

Table 6. The value of the aggregate index, its dynamics and the rank of voivodeships for the years 2015-2018 in all groups

Voivodeship	2015		2016		2017		2018		Dynamics
Małopolskie	60.7	1	59.3	1	56.8	1	64.1	1	101.8%
Mazowieckie	55.5	2	48.6	2	49.2	2	48.4	2	95.5%
Pomorskie	33.1	3	38.7	3	44.7	3	40.8	3	107.2%
Wielkopolskie	27.9	4	30.0	4	33.3	4	27.7	4	99.7%
Śląskie	67.6	1	68.4	1	67.4	1	67.4	1	99.92%
Dolnośląskie	65.1	2	64.0	2	61.3	2	66.0	2	100.5%
Kujawsko-Pomorskie	50.5	3	46.2	3	43.0	4	44.2	4	95.7%
Zachodniopomorskie	49.3	4	45.4	5	43.7	3	46.3	3	97.9%
Opolskie	47.6	5	44.7	6	41.3	6	39.6	7	94.1%
Łódzkie	46.2	6	45.5	4	42.3	5	42.7	6	97.4%
Podkarpackie	45.4	7	37.8	8	41.0	7	36.1	8	92.6%
Lubuskie	45.1	8	42.7	7	40.3	8	43.2	5	98.6%
Podlaskie	35.4	9	36.7	9	33.0	9	34.5	9	99.1%
Lubelskie	33.3	10	34.1	11	28.8	11	29.1	11	95.6%
Świętokrzyskie	32.5	11	34.3	10	32.0	10	33.8	10	101.2%
Warmińsko-Mazurskie	21.6	12	21.0	12	17.9	12	18.1	12	94.4%

Source: own calculations based on the data of Statistics Poland.

the trend is maintained, in future there should be a change to the first position of the ranking. The greatest decrease in the dynamics in group 1 was recorded in the Mazowieckie voivodeship, and in group 2 – in the Podkarpackie voivodeship.

When analysing group 1, i.e. voivodeships at a higher level of development (higher value of the index) in terms of the analysed variables, the Małopolskie voivodeship was ranked first. In the period from 2015 to 2018, each year was characterised by the same ranking positions for regions in the group 1. However, the dynamics indicate that the Mazowieckie voivodeship may fall in the ranking in favour of the Pomorskie voivodeship, unless there are clear changes in the trends for the analysed variables.

In the case of group 2, the first place in the ranking in each of the analysed years, was occupied by the Śląskie voivodeship. In the period under review, there was a change in the third position in the ranking, The Kujawsko-Pomorskie voivodeship dropped to fourth place, while the Zachodniopomorskie voivodeship moved up to third. Maintaining first place in the ranking by the Śląskie voivodeship may be difficult, but with no changes in the trends of diagnostic variables, such an exchange of places with the Dolnośląskie voivodeship may take place in future.

In the next stage the values of the diagnostic variables for all the voivodeships by groups were analysed. One diagnostic variable was modified in order to unify the

direction of impact of all the variables. The unemployment rate was considered to be a destimulant at the stage of creating the aggregate index, while in order to estimate the dynamics, it was transformed into a stimulant. All the dynamic values were calculated as the average rate of change in the analysed years. This allowed to determine the variables characterised by the greatest increase and those with the lowest increase or the greatest decrease, depending on the value of the dynamics index. Table 7 shows the variables with the highest and lowest dynamics for both groups.

Table 7. Variables with the highest and lowest dynamics in the analysed period divided into two groups

	The highest dynamics	The lowest dynamics
Group 1	Assessment of the financial situation of households – very good (%)	Population density of built-up and urbanised area (persons/km ²)
	Unemployment rate	Demographic rate per 1000 population
	Average monthly disposable income per capita	New residential buildings per 1000 population Primary market, price per m ²
Group 2	Assessment of the financial situation of households – very good (%)	Population density of built-up and urbanised area (persons/km ²)
	Unemployment rate	Demographic rate per 1000 population
	Average monthly disposable income per capita	New residential buildings per 1000 population

Source: own calculations based on the data of Statistics Poland.

The variables with the highest dynamics in both groups were the same. It can be said these had a clear impact on the current trends of optimism among the population of a given region. The following variables: a very good assessment of the material situation of the households, a low unemployment rate, and the level of disposable income had a very positive impact on the development of the region. On the other hand, this development was slowed down by such variables as population density, demographic rate per 1 000 population and the number of new completed residential buildings per 1 000 population. In group 1, there was also the price of a square metre of an apartment on the primary market. As with the constant structure of created clusters, the variables most important for the development or slowing down of the regions were very similarly divided into groups, with only occasional exceptions.

5. The COVID-19 pandemic and the economic situation in Polish voivodeships

Once the voivodeships were divided into homogeneous groups characterized by a higher and lower level of variables, it was decided to carry out a total analysis in groups of companies in the construction industry before and during the COVID-19

pandemic. The analysis was performed for the entire construction industry (section F according to Polish Classification of Business Activity 2007) as it was not possible to obtain results for housing construction only. Construction works connected with erecting buildings for the whole of Poland constituted about one-third of the value of construction and assembly production. Statistics Poland carried out a survey of the business climate, including the construction industry. By province, data were available from January 2011, i.e. from the period when the economy was recovering from the financial crisis (Jóźwiakowski, 2015; Lorek, 2016). The survey of the economic situation in manufacturing, construction, trade and services was conducted on the basis of a monthly survey. The subjective scope of the survey of the economic situation consisted of entities conducting economic activity in Poland operating in the form of: commercial companies, civil partnerships, state-owned enterprises, cooperatives and natural persons conducting economic activity classified according to PKD 2007 into individual sections. The samples for business climate surveys differed depending on the subjective scope – in manufacturing (section C) it covered 3490 units, in construction (section F) – 4000 enterprises, in trade (section G) – 7793 entities, in services (range from section H do S) – 4300 companies. The results were published in the report “Business trends in manufacturing, construction, trade and services 2000-2021”. Looking at the volume of output sold in the construction industry, the period between 2010 and 2012 was characterised by an upward trend, which means that there was some kind of rebound in this industry. It was decided to compare the results of the survey carried out by Statistics Poland for 2019, i.e. the year before the pandemic, and 2020, when the pandemic reached Poland. The average annual values, calculated for both years, were also compared with 2011, when the construction industry was emerging from the 2008-2010 financial crisis. In addition the study verified whether the formed groups responded differently or similarly to the pandemic.

The assessment of the economic situation of enterprises in 2019 was marked by a positive value of the indicator (the difference between the percentage of positive answers and the percentage of negative answers). The greatest optimism prevailed in Podlaskie and the greatest pessimism – in Dolnośląskie. In 2011, negative values of the index were recorded, which means that pessimism prevailed. The Lubuskie voivodeship achieved the lowest value of the indicator and Podkarpackie the highest, although still negative. The average percentage for all the voivodeships was -10.3 . The year of 2020, when the first cases of COVID-19 infection were officially reported in Poland, was also dominated by pessimism, only Podlaskie recorded a positive value of the indicator. The lowest recorded value concerned Lubuskie and the average for all the voivodeships was -8.99 .

When considering the indications of entrepreneurs by group, in the first group characterised by a higher level of development, the average indications in 2019 were slightly lower than the average for all the voivodeships. In Group 2, the situation was the opposite. Thus, it can be concluded that entrepreneurs in more developed regions

were less optimistic. At the same time, it is worth noting that they are also more pessimistic as the situation reversed in 2020, although the difference in indications was minimal.

Analysing the barriers limiting the activity of enterprises, clear differences in the considered periods were visible for the following:

- insufficient demand,
- uncertainty about the general economic situation,
- other.

Figure 2 presents the percentage of indications for insufficient demand.

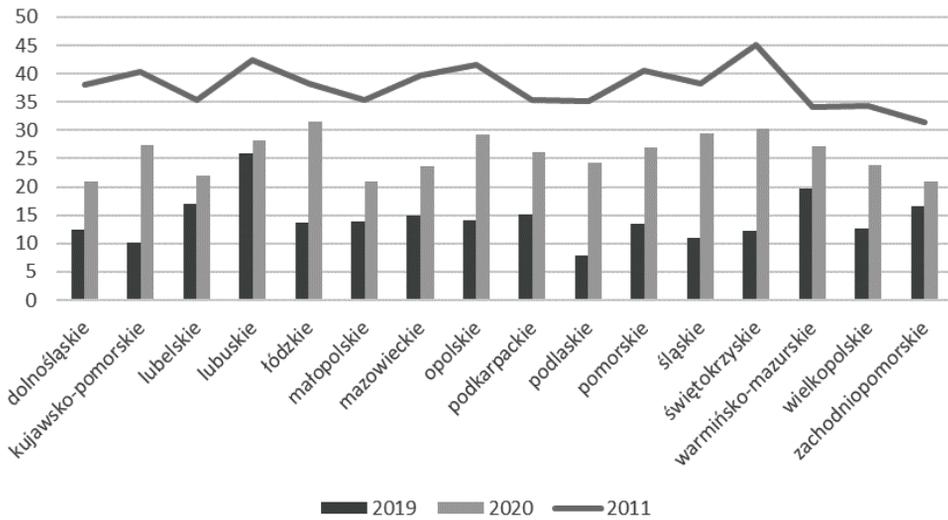


Fig. 2. The respondents' indications of insufficient demand between 2019 and 2020, and in 2011

Source: own study based on data of Statistics Poland.

The line depicting the situation in 2011 was much higher than the bars representing the indications in 2020, hence the financial crisis caused greater pessimism among entrepreneurs. The largest increase in the indicator was recorded in Śląskie and the smallest in Lubuskie, where the value of the indicator in 2019 was the highest. All the entrepreneurs identified a decline in demand as one of the issues that limited their business in 2020. By group, the more developed voivodeships recorded a lower average indicator than the less developed voivodeships. Group 1 voivodeships reported a more pessimistic assessment on the analysed barrier than Group 2, both in 2019 and 2020. A larger percentage of firms indicated problems with skills shortages, but the pandemic did not exacerbate this problem. All the voivodeships saw a decrease in the rate in comparison to 2019. Figure 3 shows the value of the indicator for that barrier – the uncertainty of the general economic situation.

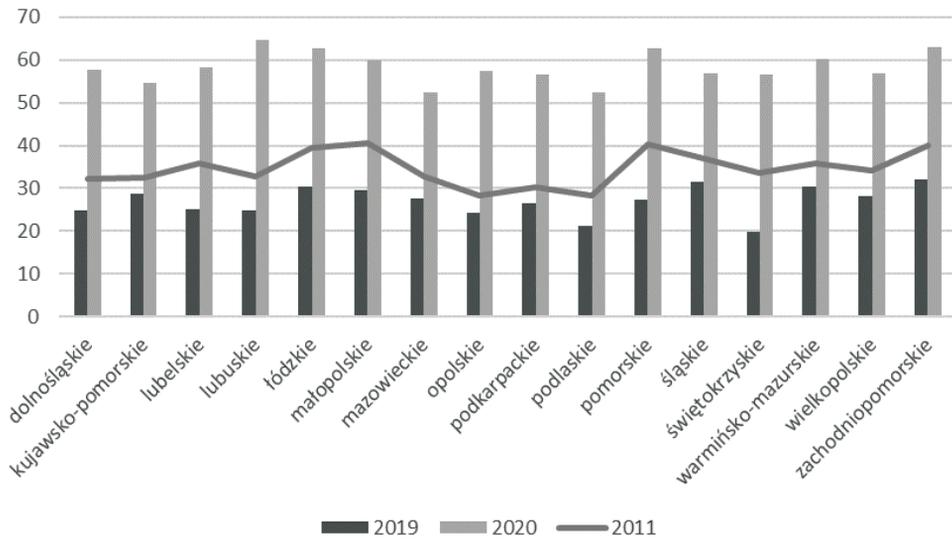


Fig. 3. The respondents' indications of uncertainty about the general economic situation between 2019 and 2020, and in 2011

Source: own study based on data of Statistics Poland.

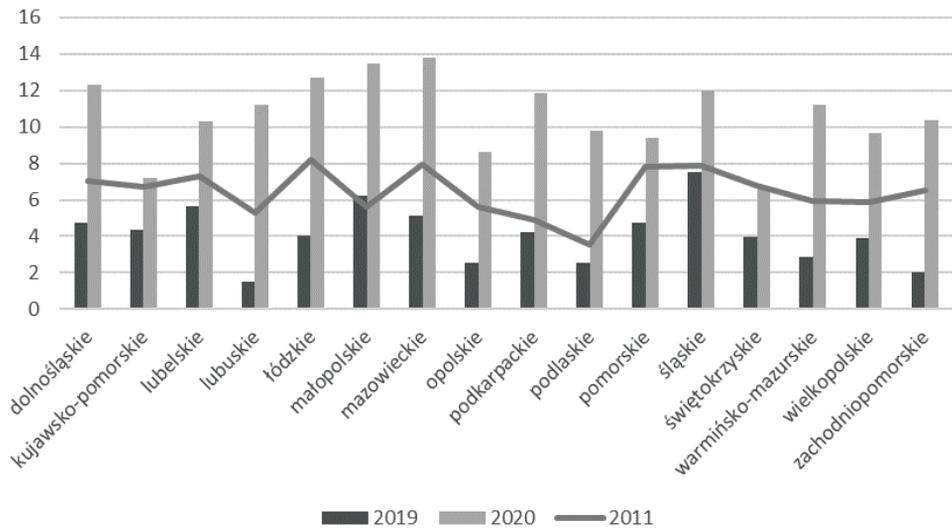


Fig. 4. The respondents' indications of barriers in the category "other" between 2019 and 2020, and in 2011

Source: own study based on data of Statistics Poland.

There was a clear increase in the percentage of companies in 2020 that indicate this barrier as one of the reasons for limiting their activities. The values shown were also significantly higher than those recorded in 2011. The highest increase in the percentage of indications for this barrier was noted in Lubuskie, and the lowest in Mazowieckie. Entrepreneurs from the Lubuskie voivodeship had the most concerns about the uncertainty related to the general economic situation during the pandemic, while those from the Mazowieckie and Podlaskie voivodeships had the least. By group, in 2019, the more developed voivodeships in Group 1 recorded an average above the average level for all the voivodeships, while those in Group 2 recorded an average below that figure. This indicates greater concern about the general economic situation among the entrepreneurs from voivodeships at a slightly higher level of development. In 2020, companies operating in the voivodeships of Group 1 indicated a lower level concerning the assessment of the general economic situation. In the Statistics Poland survey, one of the barriers that the respondents could indicate was the 'other' category, which may indicate barriers caused by the pandemic situation in Poland, therefore this category is also shown in Figure 4.

In each voivodeship, an increase in the value of the index was recorded. The largest was again in Lubuskie, and the smallest in Świętokrzyskie, while Mazowieckie was the one which in 2020 recorded the highest percentage of companies indicating this broad category of barriers limiting their activities; the lowest percentage of indications was in Świętokrzyskie. In April 2020, when the greatest jumps in the value of indicators appeared in the survey, the highest percentage of companies indicating other barriers was visible in Lubuskie, where in the following months the tendency was downward, reaching one of the lowest levels among the voivodeships at the end of that year; the lowest jump concerned Świętokrzyskie. Group 1 voivodeships reported a higher average percentage of indications for this barrier in both 2019 and 2020 than the average for all the voivodeships, which suggests that the respondents from these voivodeships indicated other activity constraints that affected them both before and during the pandemic.

The entrepreneurs also assessed the nearest future (three months) and predictions concerning the future situation of the company. Figure 5 shows the average value of the index for the respondents' answers.

In every voivodeship, there were more pessimistic assessments for the nearest future, while positive responses prevailed in the vast majority of voivodeships in 2019, with only Śląskie recording a marginally negative index value. In 2020, negative responses began to predominate, with all the indicator values being negative. The largest decrease in the indicator was recorded in the Lubuskie and Podlaskie voivodeships (the highest positive value of the indicator in 2019). The smallest decline in the index was in the Łódź voivodeship, which had one of the lower positive rates in 2019. The highest value of the indicator in 2020 was recorded in Lubuskie, and the lowest in Pomorskie. Relating the respondents' predictions about the near future to the assessment of the current situation, it can be concluded that the situation was viewed more optimistically by the companies (the average value of the indicator

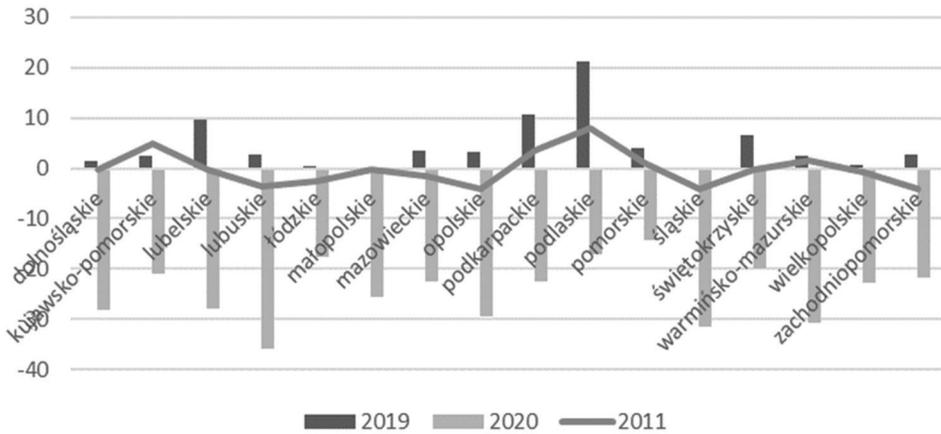


Fig. 5. The respondents' indications of the future general situation of the company between 2019 and 2020, and in 2011

Source: own study based on data of Statistics Poland.

in 2019 for the voivodeships was 8.89, and in 2020, it was -8.99). However, in the forecast the respondents on average indicated in 2019 the level of 4.45 (half the current assessment), and in 2020 at -24.3 , which was significantly lower than the current situation assessment. When considering the prediction of the future situation in both groups, the respondents from the voivodeships in group 1 in 2019 recorded an average value of the indicator lower (positive) than for all the voivodeships analysed together, and in 2020 it was higher (negative, closer to zero). This indicates that the respondents from the more developed voivodeships assessed the immediate future slightly more negatively in the year before the COVID-19 pandemic than the respondents from the less developed voivodeships, and they were more pessimistic about the future in the pandemic year. In 2020, the level of the indicator was significantly lower than in 2011, where five voivodeships had a positive value of the indicator. As regards the current and projected financial situation of enterprises in the construction sector, the results obtained are presented in Figures 6 and 7.

In 2019, the average for the indicator for all the voivodeships was -3.78 , thus there was a slight advantage for pessimistic assessments of the current situation, while the average for the expected financial situation in the near future was -0.38 . Therefore, one can conclude that the entrepreneurs were somewhat more optimistic about the near future, expecting a slight improvement. When broken down into groups, companies from the more developed voivodeships were slightly more optimistic about both the current and future financial situation. Analysing 2020, the average value of the index on the current situation of enterprises amounted to -21.7 , which was much higher than the year before; pessimism dominated in every province. The index concerning the expected future financial situation reached -25.5

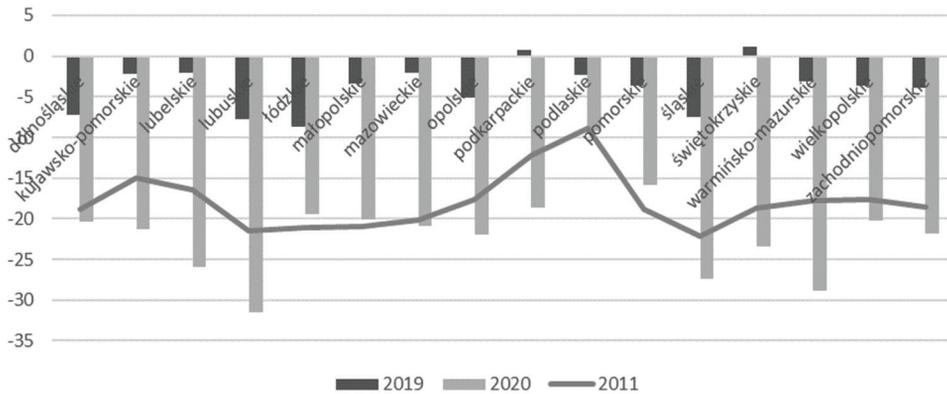


Fig. 6. Indicator values for the current financial situation of companies between 2019 and 2020, and in 2011

Source: own study based on data of Statistics Poland.

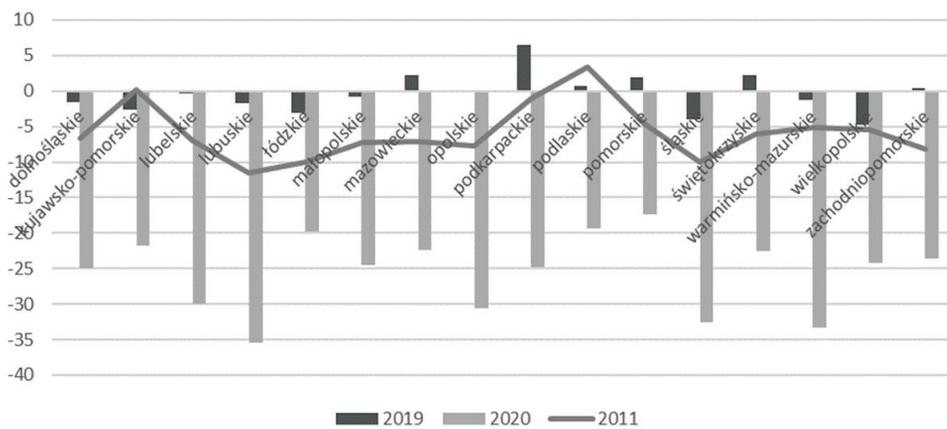


Fig. 7. Indicator values for the expected financial position of companies between 2019 and 2020, and in 2011

Source: own study based on Statistics Poland.

on average, thus the entrepreneurs saw the nearest future as slightly worse than the current situation. Considering the situation in groups, the situation looks similar to 2019. Group 1 showed greater optimism in the assessment of the current as well as the future financial situation than the average value for all voivodeships. The assessments of the entrepreneurs in 2011 were slightly more optimistic, especially regarding expectations for the near future.

Finally, the values of a composite index describing the general business climate in the construction industry were analysed. In 2019, all the voivodeships achieved a positive value of the index, so optimism (good economic situation) prevailed – the highest in Podlaskie, and the lowest in Śląskie. In 2020 the situation was reversed; all the voivodeships obtained a negative value of the indicator, so pessimism (bad economic situation) prevailed – the highest in Lubuskie, and the lowest in Podlaskie.

6. Conclusions

The basic aim of the paper, namely the division of the voivodeships into homogeneous groups, and an indication of the determinants of development in groups, was achieved. Whilst determining the groups, a hierarchical method (Ward's algorithm) was used, which proved better than the k-means method. In the case of both methods, based on Euclidean distances, the division into clusters using data from 2015 and 2018 turned out to be exactly the same, which indicates a constant structure of groups in that period. Analysing the results obtained, it can be observed that individual average values for the groups indicated a group of voivodeships with higher than average variable values and a group with lower variable values compared to the average. Comparing the obtained results with those in (Foryś, 2009; Gotowska, 2011; Kobylńska, 2016; Mach, 2014, 2019b), it can be noted that the Ward agglomeration method was most often used, but also the k-means method was found. The results obtained in the form of groups were strongly dependent on the diagnostic variables selected for the analysis. In this study, the authors used indexes as one of the methods of selecting the optimal number of clusters. In the research of other authors, the selection of the number of clusters most often resulted from the adopted criteria. The results obtained on the basis of various researchers' assumptions as well as various diagnostic variables were difficult to compare with each other because they did not have one point of reference common to all of them. In the next step, it was decided to use the iterative method of multidimensional ranking, making it possible to obtain the value of each voivodeship's potential, measured by the aggregate factor, based on the diagnostic variables. The analysis was based on previously created groups, each of which was treated as a separate set. Surprisingly, the Małopolskie voivodeship turned out to be the leader in the group of better developed voivodeships, while in the group of the less developed voivodeships, the leader was Śląskie. Looking at the average dynamics of the aggregate index for the group leaders, it can be concluded that both current leaders may lose this position in the future. The results of the analysis, which was to identify the determinants of the development of the regions, allowed for the separation of two key groups:

1. Variables with the highest dynamics:
 - Assessment of the financial situation of households – very good (%);
 - Unemployment rate;
 - Average monthly disposable income per capita.
2. Variables with the lowest dynamics:
 - Population density of built-up and urbanised area (persons/km²);

- Demographic rate per 1000 population;
- New residential buildings per 1000 population.

Group 1 was dominated by variables related to economic development, including a subjective assessment of the current situation of households. Group 2 was dominated by variables related to the demographic situation, and in addition a variable concerning the development of construction of new housing. It is worth considering a comparison of regional development strategies implemented by local governments and implementing those which give positive results by the others in the group. The comparison of individual voivodeships in groups also provided the opportunity to check the potential of the given region in terms of the analysed variables. Real estate investors can use the group leader indicators to compare the potential of another region in the same group. The search for better structured groups using other methods and distance measures seems to be justified. The response of both groups of voivodeships with a different level of real estate market development and potential, was analysed in terms of their response to the COVID-19 pandemic. Its influence on the entire construction sector was very clear, and it can be seen that the reaction of enterprises. Pessimism predominated in 2020, more than in 2019. The respondents also noted that barriers to their activities in the form of insufficient demand, the assessment of the overall economic situation, or the category of 'other' which could include emergencies such as pandemics and business restrictions, began to gain in importance. The assessment of the current situation of construction companies in 2019 was more optimistic than in 2020, which was dominated by the COVID-19 pandemic. As far as the assessment of the situation of enterprises in the near future is concerned, already 2019 showed a deterioration in sentiment (more pessimistic), but in 2020 there was a clear decline. In the future, the process of grouping the voivodeships based on the same variables, covering the period of the pandemic and post-pandemic, should be performed again. This would allow the situation of the individual groups to be assessed and compared. When broken down into groups, one can notice the reactions of companies in given regions, indicating that the level of development in which the voivodeships are located may have an impact on the more optimistic or pessimistic approach of companies to the assessment of the current situation, and the assessment of the immediate future.

References

- Alqaralleh, H., and Canepa, A. (2020). Housing market cycles in large urban areas. *Economic Modelling*, 92(C). doi.org/10.1016/j.econmod.2020.01.005
- Brzezicka, J., Łaszek, J., Olszewski, K., and Waszczuk, J. (2019). Analysis of the filtering process and the ripple effect on the primary and secondary housing market in Warsaw, Poland. *Land Use Policy*, 88. doi.org/10.1016/j.landusepol.2019.104098
- Carroquino, J., Garcia-Casarejos, N., and Gargallo, P. (2019). Classification of Spanish wineries according to their adoption of measures against climate change. *Journal of Cleaner Production*, 244, 118874. doi.org/10.1016/j.jclepro.2019.118874

- Dąbrowski, I., Mach, Ł., Mikołajczyk, Ł., and Kuświk, A. (2020). Correlation of economic development of countries with the potential of their housing real estate markets: A case study in the European Union. *European Research Studies Journal*, 23(5), 660-678.
- Foryś, I. (2009). Wykorzystanie analizy wielowymiarowej do oceny potencjału rozwoju lokalnego rynku nieruchomości mieszkaniowych. *Studia i Materiały Towarzystwa Naukowego Nieruchomości*, 17(2), 7-19.
- Foryś, I. (2011). Społeczno-gospodarcze determinanty rozwoju rynku mieszkaniowego w Polsce: ujęcie ilościowe. *Rozprawy i Studia – Uniwersytet Szczeciński*, (793), 398.
- Fritzsche, C., and Vandrei, L. (2019). The German real estate transfer tax: Evidence for single-family home transactions. *Regional Science and Urban Economics*, 74(C), 131-143. doi.org/https://doi.org/10.1016/j.regsciurbeco.2018.08.005
- Funke, M., Leiva-Leon, D., and Tsang, A. (2019). Mapping China's time-varying house price landscape. *Regional Science and Urban Economics*, 78, 103464. doi.org/10.1016/j.regsciurbeco.2019.103464
- Gotowska, M. (2011). Grupowanie województw metodą analizy skupień na podstawie transakcji kupna sprzedaży nieruchomości gruntowych. *Studia i Materiały Towarzystwa Naukowego Nieruchomości*, 19(3), 73-84.
- Huang, Y., Li, Q., Liow, K. H., and Zhou, X. (2020). Is housing the business cycle? A multiresolution analysis for OECD countries. *Journal of Housing Economics*, 101692. doi.org/10.1016/j.jhe.2020.101692
- Jóźwiakowski, P. (2015). Kryzys finansowy – przebieg i skutki dla gospodarki. *Zeszyty Naukowe Polskiego Towarzystwa Ekonomicznego w Zielonej Górze*, 2(2), 6-17.
- Kobylińska, M. (2016). Metoda aglomeracyjna w ocenie przestrzennego zróżnicowania obrotu lokalami mieszkalnymi oraz nieruchomościami i zabudowanymi budynkami mieszkalnymi. *Metody Ilościowe w Badaniach Ekonomicznych*, 17(3), 73-83.
- Kucharska-Stasiak, E. (2016). *Ekonomiczny wymiar nieruchomości*. Warszawa: Wydawnictwo Naukowe PWN.
- Leszczyński, R., and Olszewski, K. (2017). An analysis of the primary and secondary housing market in Poland: Evidence from the 17 largest cities. *Baltic Journal of Economics*, 17(2), 136-151. doi.org/10.1080/1406099X.2017.1344482
- Łaszek, J., Leszczyński, R., and Olszewski, K. (2017). The development of residential and commercial real estate, and economic development in Polish voivodeships. *Barometr Regionalny*, 15(2), 19-33.
- Łaszek, J., and Olszewski, K. (2018). Regional development of residential and commercial real estate in Poland and the risk of real estate cycles. *Barometr Regionalny*, 16(1), 41-51.
- Łaszek, J., Olszewski, K., and Waszczuk, J. (2017). Regional diversification of saturation of the housing stock. Poland against the background of Europe. *Barometr Regionalny*, 15(2), 7-17.
- Lorek, P. (2016). Wpływ kryzysu finansowego na poziom innowacyjności gospodarek krajów europejskich. *Studia Ekonomiczne. Zeszyty Naukowe Uniwersytetu Ekonomicznego w Katowicach*, (270), 172-180.
- Mach, Ł. (2014). Próba budowy homogenicznych grup województw w obszarze lokalnych rynków nieruchomości mieszkaniowych. *Metody Ilościowe w Badaniach Ekonomicznych*, 15(3), 219-227.
- Mach, Ł. (2019a). Measuring and assessing the impact of the global economic crisis on European real property market. *Journal of Business Economics and Management*, 20(6), 1189-1209. doi.org/10.3846/jbem.2019.11234
- Mach, Ł. (2019b). Zastosowanie wybranych metod WAP do oceny poziomu rozwoju lokalnych rynków nieruchomości mieszkaniowych. *Nieruchomość w przestrzeni*, 5, 135-147.
- Mach, Ł., Zmarzły, D., Dąbrowski, I., and Frącz, P. (2021). Identification and parametrization of polycyclicality in the primary housing market. *Journal of Housing and the Built Environment*. doi.org/10.1007/s10901-020-09817-6
- Malesa, K., and Malesa, T. (2018). Kredyt hipoteczny jako główne źródło finansowania inwestycji na rynku nieruchomości. *Zarządzanie Innowacyjne w Gospodarce i Biznesie*, 1(22), 140-159.

- Mikołajczyk, Ł. (2017). Podział miast wojewódzkich na jednorodne grupy pod względem ich lokalnych rynków nieruchomości przy wykorzystaniu analizy skupień. *Przegląd Nauk Stosowanych. Wybrane problemy rynku nieruchomości*, 14, 106-118.
- Milligan, G. W., and Cooper, M. C. (1987). Methodology review: Clustering methods. *Applied Psychological Measurement*, 11, 329-354. doi.org/10.1177/014662168701100401
- Milligan, G. W., and Cooper, M. C. (1988). A study of standardization of variables in cluster analysis. *Journal of Classification*, 5, 181-204. doi.org/10.1007/BF01897163
- Senetra, A., and Szarek-Iwaniuk, P. (2019). The implementation of the Wrocław taxonomic method for the identification and evaluation of problem areas in the Warmia and Mazury Region in Poland – a case study. *Socio-Economic Planning Sciences*, 67, 43-57. doi.org/10.1016/j.seps.2018.09.003
- Sokołowski, A., and Markowska, M. (2017). Iteracyjna metoda liniowego porządkowania obiektów wielocechowych. *Przegląd Statystyczny*, LXIV(2), 153-162.
- Surówka, M. (2018). Deficyt mieszkaniowy jako główny problem polskiego mieszkalnictwa. *Studia Ekonomiczne. Zeszyty Naukowe Uniwersytetu Ekonomicznego w Katowicach*, 358, 232-242.
- Walesiak, M. (2007). Procedura analizy skupień z wykorzystaniem programu komputerowego 'clusterSim' i środowiska R. *Prace Naukowe Akademii Ekonomicznej*, 15(7), 44-56.
- Walesiak, M., and Dudek, A. (2006). Symulacyjna optymalizacja wyboru procedury klasyfikacyjnej dla danego typu danych – oprogramowanie komputerowe i wyniki badań. *Klasyfikacja i analiza danych – teoria i zastosowania*, 13, 120-129.
- Walesiak, M., and Dudek, A. (2019). Package 'clusterSim'. doi.org/10.1007/BF02294245
- Walesiak, M., and Gatnar, E. (2009). *Statystyczna analiza danych z wykorzystaniem programu R*. Warszawa: Wydawnictwo Naukowe PWN.
- Wang, L., Li, S., Wang, J., and Meng, Y. (2019). Real estate bubbles in a bank-real estate loan network model integrating economic cycle and macro-prudential stress testing. *Physica A: Statistical Mechanics and its Applications*, 542(C), 122576. doi.org/10.1016/j.physa.2019.122576
- Xie, Z., Chen, S., and Wu, A. (2019). Asymmetric adjustment, non-linearity and housing price bubbles: New international evidence. *North American Journal of Economics and Finance*, 50, 101036. doi.org/10.1016/j.najef.2019.101036

IDENTYFIKACJA JEDNORODNYCH GRUP RYNKÓW NIERUCHOMOŚCI ORAZ ODPORNOŚĆ NA PANDEMIĘ SEKTORA BUDOWLANEGO

Streszczenie: W procesie badawczym wyodrębniono jednorodne grupy województw oraz determinanty rozwoju regionów. Wykorzystano metodę hierarchiczną z algorytmem Warda i metodę k -średnich dla dwóch okresów: roku 2015 i roku 2018. W celu wyodrębnienia determinantów rozwoju została zastosowana iteracyjna metoda rangowania wielowymiarowego. Uzyskane wyniki badań sugerują podział 16 województw na dwie grupy, w których wskazano liderów o największym potencjale. Utworzono listę determinantów o największej i najmniejszej dynamice. Rezultaty analizy mogą być wykorzystane przez samorządy wojewódzkie i potencjalnych inwestorów na rynku nieruchomości. Dokonano także analizy statystycznej opartej o wyniki badania koniunktury w budownictwie, głównym celem porównania jest wskazanie różnic z okresu przed pandemią COVID-19 oraz w czasie trwania pandemii. Według GUS rok 2020 zdominowali pesymiści w ocenie zarówno barier ograniczających prowadzenie działalności, jak i aktualnej i przyszłej sytuacji przedsiębiorstw.

Słowa kluczowe: analiza skupień, rynek nieruchomości, ranking, potencjał, pandemia, COVID-19.