ECONOMETRICS. EKONOMETRIA
Advances in Applied Data Analysis
Year 2019, Vol. 23, No. 4
ISSN 1507-3866; e-ISSN 2449-9994

ASSESSMENT OF CHANGES IN POPULATION AGEING OF POLISH PROVINCES IN 2002, 2010 AND 2017 USING THE HYBRID APPROACH

Marek Walesiak
Wroclaw University of Economics and Business, Wroclaw, Poland
e-mail: marek.walesiak@ue.wroc.pl
ORCID: 0000-0003-0922-2323

Grażyna Dehnel
Poznan University of Economics and Business, Poznan, Poland
e-mail: grazyna.dehnel@ue.poznan.pl
ORCID: 0000-0002-0072-9681

DOI: 10.15611/eada.2019.4.01
JEL Classification: J11, J14, C38, C43, C88

Abstract: The study investigates the variation in population ageing in Polish provinces in 2002, 2010 and 2017. Population ageing was assessed using the median age, proportion of elderly people, double ageing index, ageing index, and old-age dependency ratio. The authors took into account causes that explain changes occurring at the bottom and at the top of the population pyramid. By applying the hybrid approach combining multidimensional scaling with linear ordering (the two-step approach), the authors identified differences in the level of population ageing in a two-dimensional space. The paper applies a new method of automatic data collection from the Local Data Bank using the BDL package and the API interface (Application Programming Interface). The BDL API is a data-sharing service through webservice defining programming interfaces independent of the programming language, whereas the bdl package using this API enables webservice integration with the R statistical environment, eliminating the need for manual data extraction and enabling the automation of recurring activities.

Keywords: ageing, multidimensional scaling, aggregate measures, R program.

1. Introduction

The problems of elderly people and population ageing have been the subject of demographic and social research for many years. The demographic literature on population ageing is extensive and explores various aspects of the problem. Population ageing is analysed in the context of economic issues, living conditions, the welfare state, retirement security and the consequences for the labour market (Bongaarts, 2004; Börsch-Supan, 2003; Lee and Edwards, 2002; Lee and Mason, 2010; Metz, 2000). Another important group of topics includes questions of social assistance, health care and disability (Miller, 2001).

In the second decade of the 21st century, population ageing in European countries has reached such a high level that it is now a major demographic problem addressed by public administration institutions, social and research organisations and local government (Murphy, 2017). This shows the seriousness of the problem and calls for more in-depth research of the phenomenon.

The purpose of the study described in this article was to measure the dynamics and territorial variation of population ageing across the provinces of Poland. The empirical study was based on statistical data from the Local Data Bank about the degree of population ageing in 16 provinces for 2002, 2010 and 2017.

Differences in the degree of population ageing and its regional variation were analysed using multidimensional scaling combined with linear ordering. The analysis of the changes observed in the selected reference years was limited to the identification of the direct causes linked to fertility and mortality dynamics and the consequences of the existence of certain demographic structures. The analysis does not include indirect causes such as migration due to the lack of reliable data.

The novelty of the article consists in the application of the hybrid method (the two-step approach) to measure the dynamics and variation of population ageing in Polish provinces for two intervals: 2010-2002 and 2017-2010. The data for the study were obtained automatically by means of the **`bdl`** R package and the API of the Local Data Bank. In addition, the article provides an assessment of the stability of results of linear ordering of provinces in the reference period based on the comparison with five other aggregate measures.

The study of the dynamics and variation of population ageing across the provinces was motivated, above all, by the rapid changes in the demographic processes observed during the period of economic transformation in Poland. The long-term consequences of changes in fertility patterns and growing life expectancy are evident only after many years of observation, which is why the study spans a period of 15 years.

Territorial variation in population ageing has been investigated in many studies, since it is precisely by means of comparative analysis that one can observe differences and identify causes. The study by (Murphy, 2017) covered 11 European countries, with the exclusion of the countries of Central and Eastern Europe. Changes in demographic structures of all European countries were addressed in other studies (Káčerová, Ondačková, and Mládek, 2014; Knapik 2012). In the case of the Visegrád Group, similar studies were conducted by (Dehnel, Gołata, Obrębalski, and Walesiak, 2018; Dehnel, Gołata, and Walesiak, 2019; Káčerová and Ondačková, 2015).

Our comparative analysis is based on relative measures (cf. (Abramowska--Kmon, 2011)). In the literature of the subject there are many relative measures of ageing, so it is always possible to raise objections concerning the selected set of measures. Nonetheless, empirical studies indicate that regardless of the choice of measures, the analytic results are convergent (Murphy, 2017).

2. Research method

The study is based on data from three different reference periods. The starting point is data matrix $\mathbf{X}' = [x'_{ij}]_{3n \times m}$ (x'_{ij} – value of the *j*-th variable for *i*-th object; $j = 1, \dots m$ – variable number; m – number of variables; $i = 1, \dots, 3n$ – object number; $3n$ – number of objects including n objects in three reference periods). The objects of interest can be described in terms of preference variables, which can be divided into stimulants, destimulants and nominants. Definitions of stimulants, destimulants can be found in (Hellwig, 1981, p. 48), while that of nominants in (Borys, 1984, p. 118). These definitions are also provided in (Walesiak, 2018). Since the study aims to track changes over time, it is necessary to:

a) convert nominants into stimulants. This conversion is necessitated by the structure of the anti-pattern object;

b) find a common pattern for an anti-pattern object on the basis of the data matrix $\mathbf{X}'$. The pattern object (upper pole) represents the most favourable values of the variables, while the anti-pattern object (lower pole) – the least favourable values. In the set of objects of interest, the pattern object represents an object where population ageing is most advanced, while the anti-pattern object – an object where the process is least advanced. After accounting for the pattern and the anti-pattern object, the data matrix is denoted by $\mathbf{X} = [x_{ij}]_{(3n+2) \times m}$;

c) ensure comparability of variables by normalizing the values of the variable in the data matrix which contains combined data from the three reference periods plus the pattern and anti-pattern object. The normalized data matrix is denoted by $\mathbf{Z} = [z_{ij}]_{(3n+2) \times m}$.

The objects in the normalized data matrix were ordered in terms of the advancement of population ageing by means of the hybrid approach (a two-step procedure) proposed by (Walesiak, 2016), which makes it possible to visualize the results of linear ordering.

In the first step, multidimensional scaling is applied to visualize the arrangement of the objects in a two-dimensional space. Multidimensional scaling (MDS) is a method of mapping $f\colon [\delta_{ik}(\mathrm{Z}) \to d_{ik}[\mathrm{V}]]$ a matrix of distances between objects in an *m*-dimensional space $[\delta_{ik}(\mathrm{Z})]$ into a matrix of distances between objects in an *s*-dimensional space $d_{ik}[\mathrm{V}]$ ($s < m$; $\mathrm{V} = [v_{ij}]_{(3n+2) \times s}$ – data matrix in an s-dimensional space) to facilitate a graphic representation (visualization) and interpretation of relations between objects of interest. To visualize the results of linear ordering, s is equal to 2. The iterative procedure of multidimensional scaling using the **smacof** algorithm is described in (Borg and Groenen, 2005, pp. 204-205).

In this study, the authors used a solution which enables the selection of an optimal MDS procedure for a given normalization method, distance measure and scaling models (out of ten normalization methods, five distance measures, and four scaling models). The procedure implemented in the **mdsOpt** R package (Walesiak, Dudek, 2018b) involves the **smacofSym** function from the **smacof** package (Mair, Leeuw, Borg, and Groenen, 2018). The main idea of the **mdsOpt** package is captured in (Borg and Groenen, 2018, p. 86): "Stress is a *summative* index for *all* proximities. It does not inform the user how well a *particular* proximity value is represented in the given MDS space. (…) The least one can do is to take a look at the stress-per-point values". The optimal MDS procedure in the **mdsOpt** package (Walesiak and Dudek, 2018b) was selected using Kruskal's $Stress$-1 goodness-of-fit measure and the Herfindahl-Hirschman Index (*HHI*), calculated from stress contributions of objects (in percentages) in $Stress$-1 values (*stress per point*). Out of all the MDS procedures for which $Stress\text{-}1_p \leq critical\ stress$, the authors select the procedure that satisfies the criterion $\min_p\{HHI_p\}$ (p – id number of the MDS procedure).

In the second step, the objects are linearly ordered on the basis of the aggregate measure d_i (see (Hellwig, 1981, p. 62)):

$$d_i = \sqrt{\sum_{j=1}^{2}\left(v_{ij} - v_{+j}\right)^2} \Big/ \sqrt{\sum_{j=1}^{2}\left(v_{+j} - v_{-j}\right)^2}, \tag{1}$$

where: d_i – advancement of population ageing,
v_{ij} – j-th coordinate of i-th object in a two-dimensional MDS space,
v_{+j}, v_{-j} – j-th coordinate of the pattern and anti-pattern object in a two-dimensional MDS space.

Values of the aggregate measure d_i belong to interval $[0; 1]$. The lower the value of d_i, the younger the population of a given province. The objects of interest are arranged according to ascending values of the aggregate measure (1).

As a result of applying the optimal MDS procedure, a data matrix is obtained which is used for visualizing and interpreting the results in a two-dimensional space (the results of multidimensional scaling) and unidimensional space (the results of linear ordering based on the aggregate measure d_i).

In a diagram in a two-dimensional space (the results of multidimensional scaling), two points representing the anti-pattern and pattern object, are joined by a straight line, which forms the so-called set axis (see (Hellwig, 1981, pp. 61-62)). Isoquants of development (curves of equal development) are drawn from the pattern object. Objects located between the isoquants represent a similar level of development. The same level can be achieved by objects located at different points along the same isoquant of development (due to a different configuration of variable values). This kind of visualization enhances the interpretation of the output of linear ordering.

3. Empirical results

The empirical study draws on statistical data describing the advancement of population ageing in 16 provinces in Poland for 2002, 2010 and 2017. The selection of the measures of population ageing were mainly based on studies conducted to date (cf. (Abramowska-Kmon, 2011; Dehnel, Gołata, and Walesiak, 2019)). A set of five metric variables was finally chosen, which ensures the highest degree of differentiation between the provinces[1]:

x1 – median age (in years),
x2 – proportion of people aged 65 or older in total population (in %),
x3 – ageing index – number of people aged 65 or older per 100 persons aged 0-14,
x4 – double ageing index – proportion of people aged 80 or older in the population of people aged 65 or older (in %),
x5 – old-age dependency ratio – number of people aged 65 or older per 100 persons aged 15-64 (in %).

The above measures are classified as traditional measures of population ageing, based on chronological (calendar) age. The literature also provides alternative measures of population ageing (cf. (Sanderson, Scherbov, 2010)) which are based on remaining life expectancy (RLE)[2].

Since higher values of variables x1-x5 represent more advanced stages of population ageing, these variables should be regarded as destimulants. Data for calculations in R were obtained automatically from the Local Data Bank using the **bdl** package and the Application Programming Interface (API).

The final data matrix consisted of 50 objects, including 48 objects (data from 16 provinces for 2002, 2010 and 2017) plus a pattern object with coordinates $P = (33.57941;\ 11.031;\ 56.21;\ 14.857;\ 15.908)$ and an anti-pattern object, with coordinates $AP = (42.3297; 18.898; 133.844; 29.097;\ 28.245)$.

The optimal scaling procedure was selected from 200 MDS procedures taking into account ten normalization methods (n1, n2, n3, n5, n5a, n8, n9, n9a, n11, n12a – see (Walesiak, 2018)), five distance measures (Manhattan, Euclidean, Squared

[1] The old age threshold was set at 65 years.

[2] Current chronological (calendar) age represents the number of years a person has lived; remaining life expectancy indicates how many years the average person of a given chronological age has left to live according to the life table.

Euclidean, Chebyshev, GDM1[3] – see e.g. (Everitt, Landau, Leese, and Stahl, 2011, pp. 49-50)) and four scaling models (ratio, interval, mspline of second and third degree – (Borg, Groenen, 2005, p. 202)). As a result of applying the **mdsOpt** package, the optimal MDS procedure was selected, using the normalization method n2 (positional standardization), the interval scaling model and the GDM1 distance.

The use of MDS made it possible to map the proximities between the provinces of Poland in terms of population ageing, described in terms of five variables, into a two-dimensional space. Figure 1 shows the graphic representation of the results of multidimensional scaling of 50 objects depending on the advancement of population ageing in the provinces in Poland for 2002, 2010 and 2017. The line joining the anti-pattern object (AP) and the pattern object (P) is the so-called set axis, which can be interpreted as a path of optimal development in terms of

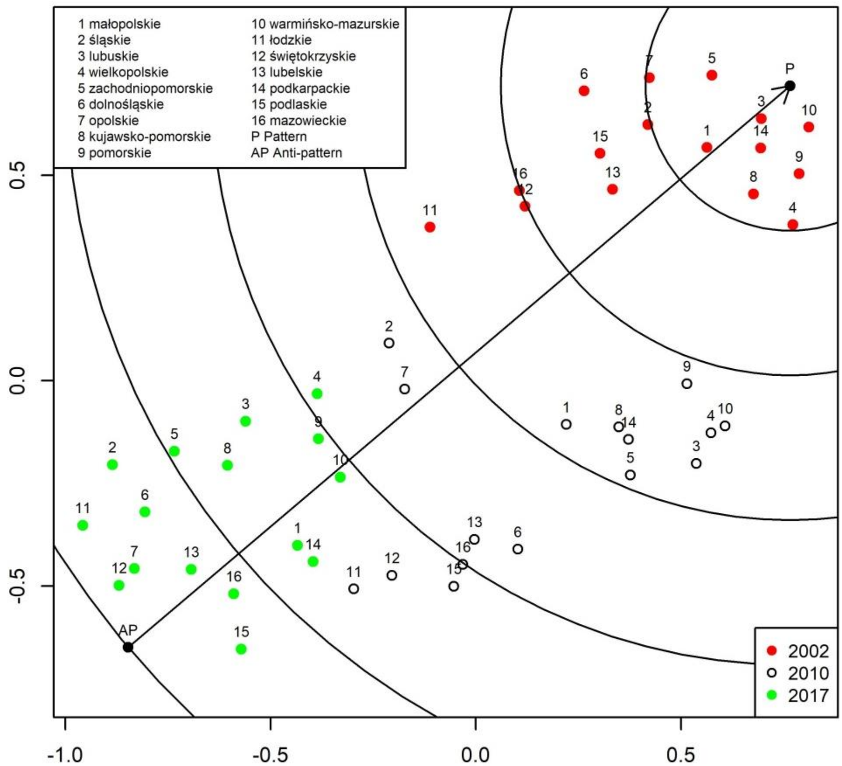

Fig. 1. Results of multidimensional scaling of 50 objects (16 provinces in 2002, 2010 and 2017; pattern and anti-pattern) depending on the advancement of population ageing

Source: diagram produced as an output of calculations in R.

[3] Cf. (Jajuga, Walesiak, and Bąk, 2003).

population ageing. Six isoquants divide the set axis into six equal parts. The further away from the pattern object a given isoquant, the older the population of the provinces which were mapped onto it.

The next step involved calculating values of the aggregate measure (1). Table 1 shows the ranking of the 50 objects depending on the advancement of population ageing according to the ascending values of measure (1) for 2002. The calculations were made using the `clusterSim` package (Walesiak and Dudek, 2018a) of the R programme (R Core Team, 2019).

Table 1. Ranking of Polish provinces depending on the advancement of population ageing (according to values of d_i) for 2002, 2010 and 2017

Province	2002		2010		2017		Rank change			Δd_i		
	d_i	R	d_i	R	d_i	R	a	b	c	a	b	c
Lubuskie	0.0502	1	0.4481	6	0.7368	4	-5	2	-3	0.3980	0.2887	0.6866
Warmińsko-mazurskie	0.0523	2	0.3988	2	0.6868	3	0	-1	-1	0.3465	0.2880	0.6345
Podkarpackie	0.0792	3	0.4479	5	0.7761	6	-2	-1	-3	0.3687	0.3282	0.6970
Zachodniopomorskie	0.0911	4	0.4845	8	0.8250	8	-4	0	-4	0.3935	0.3405	0.7339
Pomorskie	0.1018	5	0.3631	1	0.6789	2	4	-1	3	0.2613	0.3159	0.5771
Małopolskie	0.1191	6	0.4677	7	0.7759	5	-1	2	1	0.3485	0.3083	0.6568
Kujawsko-pomorskie	0.1314	7	0.4398	4	0.7819	7	3	-3	0	0.3083	0.3422	0.6505
Wielkopolskie	0.1600	8	0.4099	3	0.6501	1	5	2	7	0.2499	0.2402	0.4901
Opolskie	0.1623	9	0.5650	10	0.9381	14	-1	-4	-5	0.4027	0.3732	0.7758
Śląskie	0.1699	10	0.5490	9	0.8945	12	1	-3	-2	0.3790	0.3456	0.7246
Podlaskie	0.2324	11	0.6946	14	0.9060	13	-3	1	-2	0.4623	0.2113	0.6736
Lubelskie	0.2367	12	0.6366	12	0.8868	10	0	2	2	0.3999	0.2502	0.6501
Dolnośląskie	0.2373	13	0.6191	11	0.8909	11	2	0	2	0.3818	0.2718	0.6536
Mazowieckie	0.3343	14	0.6679	13	0.8678	9	1	4	5	0.3336	0.2000	0.5336
Świętokrzyskie	0.3354	15	0.7271	15	0.9637	16	0	-1	-1	0.3917	0.2366	0.6283
Łódzkie	0.4458	16	0.7672	16	0.9593	15	0	1	1	0.3215	0.1921	0.5135
Parameters	2002		2010		2017		X			Change in d_i		
Mean	0.1837		0.5429		0.8262		X			0.3592	0.2833	0.6425
Standard deviation	0.1095		0.1241		0.0990		X			0.0533	0.0549	0.0773
Range	0.3956		0.4041		0.3136		X			X	X	X

a – 2010-2002, b –2017-2010, c – 2017-2002, R – rank.

Source: all calculations made using R.

A comparison of the MDS results (Figure 1) with the results of the linear ordering (Table 1) of 50 objects for 2002, 2010 and 2017 reveals at least two kinds of changes:

1. In the advancement of population ageing, with three possible states:

   a) a rejuvenation of the population – objects shift towards the pattern object between the reference years (there are no regions representing this state in the results),

b) a stagnation in population ageing – objects remain in the same isoquants between the reference years (there are no regions representing this state in the results),

c) progressive population ageing – objects shift along the set axis towards the anti-pattern object between the reference years.

2. A shift of objects (provinces) relative to the set axis (with objects moving closer to or further away from the axis or crossing the axis).

The two kinds of changes are not mutually exclusive but can co-occur. In the following interpretation of empirical results, each change is discussed separately.

Taking into account Figure 1 and the values of the aggregate measure d_i (Table 1), one can assess the advancement of population ageing in each province on the basis of its position relative to the pattern (P) and the anti-pattern (AP) object. The closer a province is to the pattern object and the smaller d_i, the less advanced population ageing in that province. Conversely, the greater proximity to the anti-pattern and the higher value of the d_i measure indicate more advanced population ageing.

The empirical results clearly indicate (see Figure 1 and Table 1) that in 2017, compared to 2010, all provinces are located further away from the anti-pattern (AP). The mean change in the aggregate measure is also higher $\left(\Delta\bar{d}_i = 0.2833\right)$. A similar change can be observed between 2010 and 2002 $\left(\Delta\bar{d}_i = 0.3592\right)$. Between 2002 and 2017, the mean value of the aggregate measure increased by 0.6425. This change represents a shift in the age structure of the populations in the provinces toward the older ages. The biggest change in the values of the aggregate measure in 2017 compared to 2002 (Δd_i) can be observed for opolskie (0.7758), zachodniopomorskie (0.7339), śląskie (0.7246), and podkarpackie (0.6970). The lowest change in values of the aggregate measure can be noted for wielkopolskie (0.4901), łódzkie (0.5135) and mazowieckie (0.5336).

Based on the values of the d_i measure it can be concluded that the relatively best situation in terms of population ageing in 2002 was in lubuskie (3), warmińsko-mazurskie (10) and podkarpackie (14), while the provinces most affected by the phenomenon included łódzkie (11), świętokrzyskie (12) and mazowieckie (16). There is a considerable change in the ranking of provinces according to the values of d_i between 2002 and 2017 (the Kendall rank correlation coefficient is equal to 0.5167). The biggest increase can be observed for wielkopolskie (a shift from 7 to 1), mazowieckie (from 14 to 9) and pomorskie (from 5 to 2), while the biggest decrease can be observed for opolskie (a shift from 9 to 14), zachodniopomorskie (from 4 to 8), lubuskie (from 1 to 4) and podkarpackie (from 3 to 6). In the first period (2002-2010), there was an increase, while in the second (2010-2017) – a decline in the differentiation of provinces in terms of population ageing (see the dispersion measures in Table 1).

The advancement of population ageing across the provinces in Poland in 2002, 2010 and 2017 is shown in Figure 2.

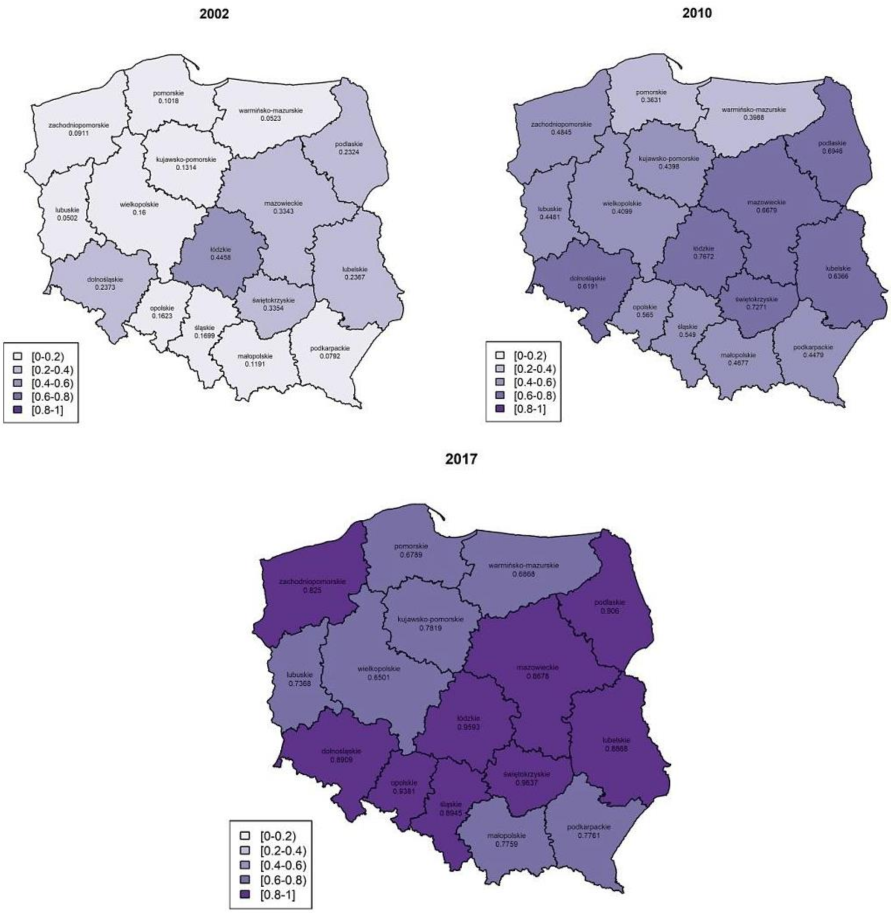

Fig. 2. Classes of provinces depending on the advancement of population ageing in 2002, 2010 and 2017 (in terms of d_i values)

Source: diagram produced as an output of calculations in R.

The class intervals represent different levels in the advancement of population ageing, from the lowest $[0-0.2]$ to the highest $[0.8-1]$. Based on Figures 1 and 2, it is possible to distinguish groups of provinces where the process of population ageing between pairs of references years (2010-2002 and 2017-2010) proceeded in a similar fashion. These are: śląskie (2) and opolskie (7); lubuskie (3), wielkopolskie (4) and kujawsko-pomorskie (8); pomorskie (9) and warmińsko-mazurskie (10); małopolskie (1) and podkarpackie (14).

The second kind of changes in the position of objects involves a shift relative to the set axis. The close proximity of provinces to the set axis (coordinates of objects located at the intersection of the isoquant with the set axis represents a situation when the normalized values of the five variables for that province are quite similar.

This is exemplified by lubuskie (3) and małopolskie (1) in 2002 and warmińsko--mazurskie (10) in 2017. The further away from the set axis a province is located, the more dissimilar the normalized values of the diagnostic variables are.

Table 2. Shifts in the position of provinces in the two-dimensional diagram representing changes in the advancement of population ageing between 2002, 2010 and 2017

No	Province	Euclidean distance from set axis			Distance difference (*D*)			Sum of distances* (*S*)		
		2002	2010	2017	a	b	c	a	b	c
1	małopolskie	0.0163	0.2764	0.0772	0.2601	-0.1992	0.0609	0.2927		0.0935
2	śląskie	0.1526	0.1539	0.3644	0.0013	0.2105	0.2118			
3	lubuskie	0.0156	0.5534	0.2351	0.5377	-0.3183	0.2195		0.3536	0.2507
4	wielkopolskie	0.2625	0.5196	0.1736	0.2571	-0.3460	-0.0889		0.5184	0.4362
5	zachodniopomorskie	0.1431	0.4710	0.2916	0.3279	-0.1794	0.1485	0.6141	0.7885	
6	dolnośląskie	0.3150	0.4315	0.2251	0.1165	-0.2064	-0.0900	0.7465	0.6933	
7	opolskie	0.2366	0.0438	0.1366	-0.1928	0.0928	-0.1000			
8	kujawsko-pomorskie	0.1433	0.3636	0.1814	0.2203	-0.1822	0.0381		0.6566	0.3247
9	pomorskie	0.1776	0.3909	0.0877	0.2133	-0.3032	-0.0899		0.1804	0.2653
10	warmińsko-mazurskie	0.1065	0.5291	0.0181	0.4227	-0.5110	-0.0883			
11	łódzkie	0.3051	0.2468	0.2984	-0.0583	0.0516	-0.0067	0.5519	0.4786	
12	świętokrzyskie	0.1939	0.2817	0.1296	0.0878	-0.1520	-0.0642	0.4755	0.5473	
13	lubelskie	0.0874	0.3451	0.0458	0.2577	-0.2992	-0.0415	0.4324	0.5452	
14	podkarpackie	0.0693	0.4025	0.1320	0.3332	-0.2705	0.0627			
15	podlaskie	0.1739	0.3995	0.1812	0.2256	-0.2184	0.0073	0.5734		0.3550
16	mazowieckie	0.2317	0.3733	0.0666	0.1416	-0.3067	-0.1651	0.6050		0.2982

a – 2010-2002, b – 2017-2010, c – 2017-2002; *in cases when an object crosses the set axis.

Source: all calculations made using R.

To provide a more detailed description of the shifts in the position of the provinces in the two-dimensional diagram (Figure 1), three additional quantities were calculated for each object: its distance from the set axis in each reference year, the difference between distances recorded for a pair of reference years and the sum of these distances (see Table 2).

If the difference *D* is positive, a given object has shifted away from the set axis. This in turn represents a greater level of dissimilarity in the normalized values of the diagnostic variables describing the advancement of population ageing in a given province. Such positive changes between the normalized values for 2010 and 2002 can be observed in 14 provinces. The biggest increase can be seen in lubuskie ($D = 0.5377$).

The negative value of *D* denotes a shift towards the set axis, or a greater similarity of values of the diagnostic variables. Changes of this kind can be observed for the majority of provinces in 2017 compared to the corresponding values for 2010 (13 objects) and for 2002 (9 objects). The biggest decline in the degree of dissimilarity between the values of diagnostic variables can be observed in warmińsko-mazurskie ($D = -0.5110$).

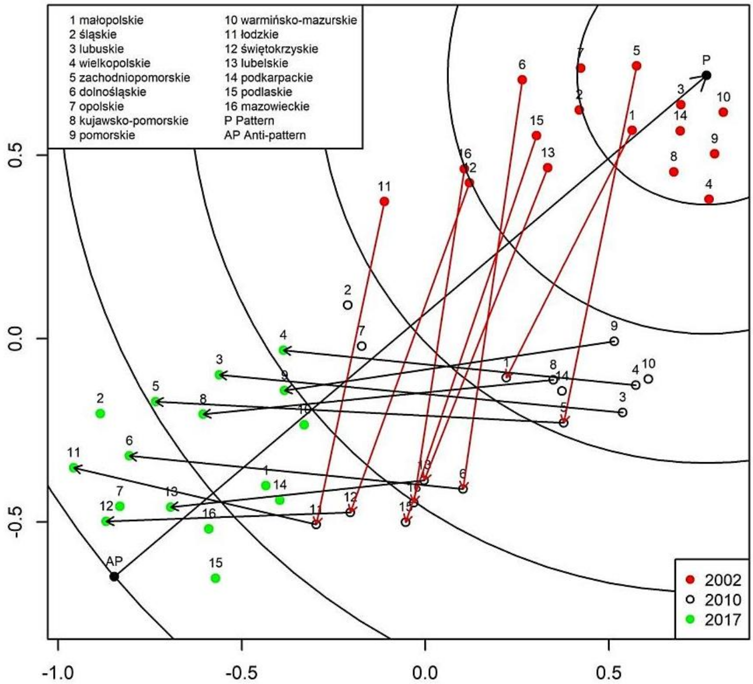

Fig. 3. Shifts of provinces across the set axis in a two-dimensional space in the period 2010-2002 (red lines) and 2017-2010 (black lines)

Source: diagram produced as an output of calculations in R.

The third group of objects identified in the analysis of the results includes those whose positions between two reference years (2010-2002, 2017-2010 and 2017-2002) shift across the set axis (see Figure 3). For each pair of such provinces, their distances from the set axis were added (S – see Table 2).

Five of the 16 provinces crossed the set axis (in 2010 and 2017), with the biggest shifts observed for two provinces: zachodniopomorskie (5) and dolnośląskie (6). Changes in the configuration of the diagnostic variables are additionally accompanied by big shifts along the set axis towards the anti-pattern object (AP), which means a considerable advancement of the ageing process (see Table 1, column Δd_i). A different situation can be observed in łódzkie (11), which is characterised by a relatively high degree of advancement of population ageing throughout the reference period. This is an example of a province which crossed the set axis twice, in each case moving between positions at a considerable distance from the set axis, but undergoing a relatively small shift along the axis towards the anti-pattern (AP) (see Table 1, column Δd_i).

The analysis of the two-dimensional diagram (Figure 3) reveal one more relationship between the objects. There are provinces with the same degree of advancement of population ageing but with a different configuration of values of the diagnostic variables (objects located at different points along the same isoquant of development). Examples of such provinces in 2002 include wielkopolskie (4), śląskie (2) and opolskie (7), which are classified as having relatively young populations. This pattern can be identified for objects from different years, as is exemplified by objects located in the vicinity of the fourth isoquant from the pattern object. Included in this are two pairs of provinces with similar degrees of population ageing in different reference years: pomorskie (9) and mazowieckie (16), and podlaskie (15) and warmińsko-mazurskie (10). Such a location means that in 2017 pomorskie (9) represented a similar degree of population ageing to that of mazowieckie (16) in 2010, whereas warmińsko-mazurskie (10) in 2017 represented a similar degree of population ageing to that of podlaskie (15) in 2010.

4. Assessment of the stability of the results of linear ordering

The stability of the results of linear ordering was assessed by comparing the results obtained by means of the two-step procedure (section 3) with those obtained using five aggregate measures (composite indicators): the measure of development (Hellwig, 1968, 1972), the TOPSIS measure (Hwang and Yoon, 1981), the GDM1 distance measure (Walesiak, 2002), the TOPSIS with GDM1 (Walesiak, 2014), and the arithmetic mean. In all cases the variables were normalized using n2 positional standardization. The results of six methods of linear ordering were compared on the basis of Kendall's tau coefficient, producing the matrix shown in Table 3.

The degree of similarity of the rankings of 48 objects (16 provinces in 2002, 2010 and 2017) with respect to the advancement of population ageing is high. The ranking of provinces based on the two-step approach are the most similar to the ranking obtained by means of Hellwig's measure. The values of Kendall's tau coefficient range from 0.9308 to 0.9875. This means that the results of linear ordering (rankings) are stable and are not significantly dependent on the method of linear ordering used. The advantage of the two-step procedure is that its results can be visualized in a two--dimensional space, which facilitates interpretation. The classical methods of linear ordering based on aggregate measures do not provide such a possibility.

Table 3. Matrix of Kendall's tau coefficients measuring correlation between values of six methods of linear ordering

SMR	Two-step approach	TOPSIS	Hellwig	GDM1	GDM1_TOPSIS	Mean
Two-step approach	1	0.9592	0.9716	0.9645	0.9468	0.9343
TOPSIS	0.9592	1	0.9734	0.9698	0.9875	0.9574
Hellwig	0.9716	0.9734	1	0.9751	0.9609	0.9485
GDM1	0.9645	0.9698	0.9751	1	0.9645	0.9308
GDM1_TOPSIS	0.9468	0.9875	0.9609	0.9645	1	0.9485
Mean	0.9343	0.9574	0.9485	0.9308	0.9485	1

Source: own calculations made using R.

5. Summary and conclusions

The measurement of population ageing is a complex process and requires a multivariate approach. To measure the advancement of the ageing process of the population of Poland in 2017 compared to 2002, the authors used the method of multidimensional scaling combined with linear ordering (the hybrid approach). The analysis was conducted at the level of provinces on the basis of five diagnostic variables. The variables were selected to account for changes occurring at the bottom and at the top of the population pyramid. Due to the lack of reliable data the analysis did not take into account indirect causes of population ageing such as migration. To track the development of the process in the intermediate period, data for 2010 were also included in the analysis.

The results clearly indicate that populations of all the provinces are getting older, which is reflected by all the objects shifting towards the anti-pattern object between the reference years. Based on the visualization of the results and the increasing values of measure d_{i} one can notice that the ageing process was more dynamic between 2002 and 2010 than between 2010 and 2017. The smallest changes were observed in wielkopolskie, łódzkie and mazowieckie. The biggest changes could be noted in opolskie, zachodniopomorskie, śląskie and podkarpackie. The results also indicate the decreasing degree of dissimilarity between provinces in terms of population ageing.

The use of multidimensional scaling leads to a partial loss of information. In this case, the configuration of objects in the original five-dimensional space was mapped into a two-dimensional space. As a result, however, it was possible to visualize the results of a multivariate phenomenon in a two-dimensional space, which provided additional scope for interpretation.

The authors are well aware of the limitations of the study that are due to the selected set of variables.

All the calculations and diagrams were prepared using scripts in the R environment.

Acknowledgements

The project is financed by the Polish National Science Centre DEC-2015/17/B/HS4/00905.

Bibliography

Abramowska-Kmon, A. (2011). O nowych miarach zaawansowania procesu starzenia się ludności. *Studia Demograficzne*, *1*(159), 3-22.

Bongaarts, J. (2004). Population aging and the rising cost of public pensions. *Population and Development Review*, *30*(1), 1-23. https://doi:10.1111/j.1728-4457.2004.00001.x

Borg, I., and Groenen, P. J. F. (2005). *Modern multidimensional scaling. Theory and applications,* 2nd Edition. New York: Springer Science+Business Media. https://doi: 10.1007/0-387-28981-X

Börsch-Supan, A. (2003). *Labor market effects of population aging*. *Labour*, *17*(s1), 5-44. https://doi.org/10.1111/1467-9914.17.specialissue.2

Borys, T. (1984). Kategoria jakości w statystycznej analizie porównawczej [The category of quality in statistical comparative analysis]. *Research Papers of University of Economics in Wrocław*, (284).

Dehnel, G., Gołata, E., and Walesiak, M. (2019). Assessment of Changes in Population Aging in Regions of the V4 Countries with Application of Multidimensional Scaling, *Argumenta Oeconomica* (accepted for print).

Dehnel, G., Gołata, E., Obrębalski M., and Walesiak, M. (2018). Ocena zmian w procesie starzenia się ludności w regionach krajów Grupy Wyszehradzkiej z zastosowaniem wybranych metod statystycznej analizy wielowymiarowej, *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, (507), 39-52. http://dx.doi.org/10.15611/pn.2018.507.04

Everitt, B. S., Landau, S., Leese, M., and Stahl, D. (2011). *Cluster analysis*, Chichester: John Wiley & Sons. https://dx.doi.org/10.1002/9780470977811

Hellwig, Z. (1968). Zastosowanie metody taksonomicznej do typologicznego podziału krajów ze względu na poziom ich rozwoju i strukturę wykwalifikowanych kadr [Procedure of evaluating high level manpower data and typology of countries by means of the taxonomic method]. *Przegląd Statystyczny*, 15 (4), 307-327.

Hellwig, Z. (1972). Procedure of evaluating high-level manpower data and typology of countries by means of the taxonomic method, In Z. Gostkowski (ed.), *Towards a system of human resources indicators for less developed countries* (pp. 115-134), Papers Prepared for the UNESCO Research Project. Wrocław: Ossolineum, The Polish Academy of Sciences Press,.

Hellwig, Z. (1981). Wielowymiarowa analiza porównawcza i jej zastosowanie w badaniach wielocechowych obiektów gospodarczych [Multivariate Comparative Analysis and its Application in Multifeature Economic Objects], In W. Welfe (Ed.), *Metody i modele ekonomiczno-matematyczne w doskonaleniu zarządzania gospodarką socjalistyczną [Economic and mathematical methods and models in the improvement of socialist economy management]* (pp. 46-68). Warsaw: PWE.

Hwang, C. L., and Yoon, K. (1981). *Multiple attribute decision making – methods and applications. A state-of-the-art survey*, New York: Springer-Verlag. ISBN: 978-3-540-10558-9. http://dx.doi.org/10.1007/978-3-642-48318-9

Jajuga, K., Walesiak M., and Bąk, A. (2003). On the general distance measure. In M. Schwaiger, O. Opitz (Eds.), *Exploratory data analysis in empirical research* (pp. 104-109). Berlin, Heidelberg: Springer-Verlag. http://dx.doi.org/10.1007/978-3-642-55721-7_12

Káčerová, M., and Ondačková, J. (2015). The process of population ageing in countries of the Visegrad Group (V4). *Erdkunde*, *69*(1), 49-68. https://doi.org/10.3112/erdkunde.2015.01.04

Káčerová, M., Ondačková, J., and Mládek, J. (2014). Time-space differences of population ageing in Europe. *Hungarian Geographical Bulletin*, *63*(2), 177-199. https://doi.org/10.15201/hungeobull.63.2.4

Knapik, O. (2012). Proces starzenia się ludności w krajach europejskich. Ujęcie historyczne i perspektywy przemian do 2050 roku. In J. Kurkiewicz (Ed.), *Demograficzne uwarunkowania i wybrane społeczno-ekonomiczne konsekwencje starzenia się ludności w krajach europejskich* (pp. 47-92). Kraków: Wydawnictwo Uniwersytetu Ekonomicznego w Krakowie. ISBN 978-83--7252-606-9.

Lee, R., and Edwards, R. (2002). The fiscal effects of population aging in the U.S.: Assessing the uncertainties. In J. M. Poterba (Ed.), *Tax policy and the economy* (pp. 141-180), MIT Press (Vol. 16). Retrieved from http://www.nber.org/chapters/c10865.pdf

Lee, R., and Mason, A. (2010). Fertility, human capital, and economic growth over the demographic transition. *European Journal of Population*, *26*(2), 159-182. http://dx.doi.org/10.1007/s10680-009-9186-x

Mair, P., De Leeuw, J., Borg, I., and Groenen, P. J. F. (2018). *Smacof: Multidimensional scaling. R package version 1.10-8*. Retrived from http://CRAN.R-project.org/package=smacof

Metz, D. H. (2000). Mobility of older people and their quality of life. *Transport Policy*, *7*(2), 149-152. https://doi.org/10.1016/S0967-070X(00)00004-4

Miller, T. (2001). Increasing longevity and medicare expenditures. *Demography*, *38*(2), 215-226. https://dx.doi.org/10.1353/dem.2001.0018

Murphy, M. (2017). Demographic determinants of population aging in Europe since 1850. *Population and Development Review*, *43*(2), 257-283. https://doi.org/10.1111/padr.12073

R Core Team. (2019). *R: A language and environment for statistical computing, R foundation for statistical computing*, Vienna. Retrived from: http://www.R-project.org

Sanderson, W.C., Scherbov, S. (2010). Remeasuring ageing. *Science*, 329(5997), 10 September, 1287-1288. https://doi.org/10.1126/science.1193647

Walesiak, M. (2002). *Uogólniona miara odległości w statystycznej analizie wielowymiarowej [The generalized distance measure in multivariate statistical analysis]*. Wrocław: Wydawnictwo Akademii Ekonomicznej we Wrocławiu.

Walesiak, M. (2014). Wzmacnianie skali pomiaru w statystycznej analizie wielowymiarowej [Reinforcing measurement scale for ordinal data in multivariate statistical analysis]. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, (327), 60-68.

Walesiak, M. (2016). Visualization of linear ordering results for metric data with the application of multidimensional scaling. *Ekonometria*, 2(52), 9-21. http://dx.doi.org/10.15611/ekt.2016.2.01

Walesiak, M. (2018). The choice of normalization method and rankings of the set of objects based on composite indicator values. *Statistics in Transition New Series*, 19(4), 693-710. https://doi.org/10.21307/stattrans-2018-036

Walesiak, M., and Dudek, A. (2018a). *clusterSim: Searching for optimal clustering procedure for a data set. R package version 0.47-3*. http://CRAN.R-project.org/package=clusterSim

Walesiak, M., and Dudek, A. (2018b). *mdsOpt: Searching for optimal MDS procedure for metric and interval-valued data. R package version 0.3-3*. http://CRAN.R-project.org/package=mdsOpt

OCENA ZMIAN W PROCESIE STARZENIA SIĘ LUDNOŚCI WOJEWÓDZTW POLSKI DLA LAT 2002, 2010 I 2017 Z WYKORZYSTANIEM PODEJŚCIA HYBRYDOWEGO

Streszczenie: Przedmiotem badania jest problem zróżnicowania starzenia się ludności województw Polski w latach 2002, 2010 i 2017. Porównanie stopnia zaawansowania procesu starzenia się ludności przeprowadzono z wykorzystaniem zmiennych, takich jak: mediana wieku, współczynnik starości, wskaźnik podwójnego starzenia, indeks starości, współczynnik obciążenia osobami starszymi. Zestaw zmiennych uwzględnia zmiany w procesie starzenia zarówno od podstawy, jak i od góry piramidy wieku. W celu porównania stopnia zaawansowania oraz zróżnicowania procesu starzenia się ludności zastosowano podejście hybrydowe łączące skalowanie wielowymiarowe z porządkowaniem liniowym, pozwalające na prezentację wyników w przestrzeni dwuwymiarowej. W artykule wykorzystano nowy sposób automatycznego pozyskiwania danych z Banku Danych Lokalnych z wykorzystaniem pakietu BDL oraz interfejsu API (*Application Programming Interface*). API BDL to usługa udostępniania danych poprzez *webservice* definiujący interfejsy programistyczne niezależne od języka programowania, natomiast pakiet BDL korzystający z API umożliwia integrację *webservice* ze środowiskiem statystycznym R, eliminując konieczność ręcznego pobierania danych oraz umożliwiając automatyzację działań cyklicznych.

Słowa kluczowe: starość, skalowanie wielowymiarowe, miary agregatowe, program R.