QUANTITATIVE METHODS IN ECONOMICS
Vol. XIV, No. 1, 2013, pp. 37 – 52

SYNTHETIC RADAR MEASURES THAT FACTOR IN WEIGHTS OF FEATURES

Zbigniew Binderman
Jacob of Paradyz State Higher Vocational School in Gorzow Wielkopolski
Bolesław Borkowski
Department of Econometrics and Statistics
Warsaw University of Life Sciences – SGGW
Wiesław Szczesny
Department of Informatics, Warsaw University of Life Sciences – SGGW
e-mail: zbigniew_binderman@sggw.pl, boleslaw_borkowski@sggw.pl,
wieslaw_szczesny@sggw.pl

Abstract: This work is a direct continuation of previous works of authors ([Binderman, Borkowski, Szczesny 2008, Binderman, Szczesny 2009]). The goal of this work is to generalize known methods of constructing measures of synthetic objects, methods that use radar charts of vectors. The methods listed in this work factor in weights assigned to features and make the values of indicators independent from the ordering of considered features. The listed methods have been applied in the analysis of regional differentiation of agriculture in Poland between 2001 and 2010.

Keywords: synthetic measure, classification, ordering, weights, radar measure.

INTRODUCTION

Multidimensional analysis of empirical material requires many specialist tools. It is necessary to not only present, in a form of a clear and understandable chart, values of many measures assessing fragmentary aspects but also produce a summary valuation of area in a form of a normalized numerical coefficient. In simple situation analysis can be aided by a spreadsheet, as it provides basic functionality to create charts and calculations. Some of smaller companies decide to adopt such solutions. However, in practice that solution proves insufficient, because normally, one needs to carry out multi-criteria analysis that require processing gathered data in multiple ways. Many applications on the market have been designed for that purpose, as they offer management of data (protecting it from loss of integrity and allowing easy access to historical data arranged by present organizational structure) and standardization of measures and their presentation across the whole company. It is especially important when one is dealing with multilevel structures of measures and multilevel organizational structures. There are many provides of such tools and applications. Large corporations choose tools of multinational providers. The global market shares of particular global providers are available in annual IDC reports [see Vesset, McDonough, Wardley, Schubmehl 2012]. In Poland some of that also have a large market share as indicated by, e.g. **[**Odrzygóźdź, Karwański, Szczesny 2009**]**. While smaller companies utilize dedicated solution prepared by local provides. An interesting one is, for example, the ProRATE application offered by PRONOST ltd.

However, the key element to success when building such information systems is the right choice of means to present the results of observed fragmentary measures as well as normalized synthetic coefficient (or a set thereof). In subsequent analysis we will concentrate only on functionality available in a spreadsheet. With such constraints in place, we believe that the best way to present a set of fragmentary measures in a given moment in time is to utilize radar charts, as they allow to present in a clear way the values of desired measures (e.g. average, the positive and negative sample, etc.). Many analysts consider the area of a polygon defined by the values of fragmentary measures as a natural summary value (a synthetic measure) when the fragmentary measures are to be treated equally. In majority of cases the value of such a synthetic measures is dependent on the ordering of fragmentary measures (see **[**Binderman, Borkowski, Szczesny 2008, Binderman, Szczesny 2009**]**).

The goal of this work is to generalized known methods of constructing measures of synthetic objects, methods that utilize radar charts. Methods listed in this work take into consideration weights assigned to features as well as make values of synthetic coefficients independent from the order of features. In general, when constructing a synthetic coefficient one is more often in a situation when fragmentary measures are not uniformly important. Most often that is the case when e.g. constructing bonus indexes in corporations with many branches, because the evaluated activities are not always equally important. This means that the user should not only be presented on the chart (in this case a radar chart) with the values of specific fragmentary measures but also their weights in the summary synthetic value. In the next section we shall present how to achieve that.

DATA VISUALISATION WITH WEIGHTS OF VARIABLES

Let us consider a typical problem of ordering m∈*N* objects: **Q**1, **Q2**,…,**Q**m defined by a set of $n \in$ N variables (features): X_1, X_2,...,X_n. Let vector $\mathbf{x_i} = (x_{i1}, x_{i2}, \ldots, x_{in}) \in \Re^n$, $n \in \mathbb{N}$, *i=1,2,...,m*, define an i-th object $\mathbf{Q}_i$.

An important phase in a comparative analysis is the calculation of synthetic coefficients which take into account non-uniform weights of variables (fragmentary coefficients). This phase is called the weighting of variables and its aim is to assign different importance to different variables. It is used because in the considered problems, variables often have different roles. This problem is yet to be fully analyzed, as literature presents different approaches and opinions as to should variables be weighted at all and if yes, how it should be done. Defining the weights of variables is tightly coupled with the ranking of information value of features, which in turn is considered when choosing diagnostic variables. That means that when choosing and weighting variables one should base their actions on knowledge of the analyzed process while statistical tools, which use information about distributions of variables, should only provide aid when necessary. Let us reiterate that one of the simplest ways of determining values of weights when aided by statistical tools is based on the coefficient of volatility of features after they have been normalized, when the assumption that they are all positive hold true:

$$w_j = \frac{V_j}{\sum_{k=1}^{n} V_k}, \quad j = 1,2...,n; \quad V_j = \frac{S_j}{\bar{x}_j},$$

where $\bar{x}_j$ – arithmetic average, S_j – standard deviation of j-th variable

Another way of determining weights utilizes a matrix of coefficients of linear correlation of diagnostic variables. The system of weights is created on the basis of measures of explicitness of variables. The explicitness of features means the certainty of classification of variables into sets of 'stimulator' and 'de-stimulator'. Weighting factors of features (assuming the number of features after reduction is equal to n) are calculated as follows [Grabiński, Wydymus, Zeliaś 1989, p.26, Zeliaś 2000]:

$$w_j = \frac{\left|\sum_{i=1}^{n} r_{ij}\right|}{\sum_{l=1}^{n}\left|\sum_{i=1}^{n} r_{il}\right|} \quad (j=1,2...,n),$$

where r_{ij} – coefficient of linear correlation between i-th and j-th variables.
If the identification is based on factor analysis then weighting factors are determined by:

$$w_j = \frac{|q_{1j}|}{\sum_{j=1}^{n}|q_{1j}|},\ j=1,...,n;$$

where q_{1j} – coefficient of linear correlation between j-th diagnostic variable and the first main factor.

Factors $q_{1j} = \sqrt{\lambda_1 a_{1j}}$, where λ_1 – maximal eigenvalue of correlation matrix $\mathbf{R} = [r_{ij}]_{n\times n}$, a_{1j} – subsequent elements of eigenvector, corresponding to value λ_1, $j = 1,2...,n$. The described above ways of determining differentiated weights fulfil the required criteria, they are non-negative: $w_j \geq 0$ and they sum to one: $\sum_{j=1}^{n} w_j = 1$

As mentioned above, the choice of a system of weights is for the researcher to make. However, when researching economic phenomenon, which can be very intricate, a demand raises to, for the purpose of analyzing statistic features of variables, consider only extra-statistic information about the process that comes from "field experts". More information on choosing a weight system can be found in the following works [Kukuła 2000, Nowak 1990, Pociecha 1996, Grabiński 1984, Borys 1984, Rusnak, Siedlecka, Siedlecki 1982, Stobiecka 1998, Młodak 2006, Malina 2004, Zeliaś 2000].

Let us introduce basic concepts and designations necessary in this work.

Definition 1. Set

$$\Omega := \left\{ w = (w_1, w_2, \ldots, w_n) \in \Re^n : \sum_{i=1}^{n} w_i = 1,\ w_i > 0 \right\},$$

will be called a set of weights and its elements Ω - vectors of weights.

Naturally, each weight is a structure, while the reverse do not always holds.

Definition 2. Let vector $\mathbf{w} = (w_1, w_2, \ldots, w_n) \in \Omega$,. Operator $\mathsf{B}_{\mathbf{w}} : \Re^n \to \Re^n$, is defined by:

$$\mathsf{B}_{\mathbf{w}}\mathbf{x} = \left(\frac{w_1 x_1}{w_{max}}, \frac{w_2 x_2}{w_{max}}, \ldots, \frac{w_n x_n}{w_{max}} \right), \tag{1}$$

Where $\mathbf{x} = (x_1, x_2, ..., x_n) \in \Re^n$, $w_{max} = \max(w_1, w_2, ..., w_n)$, will be called a weight operator induced by weight vector **w.**

Without loss of generality let us assume that features of object $\mathbf{x} = (x_1, x_2, ..., x_n) \in \Re^n$ are of stimulator variety and their values belong to the interval <0,1>, i.e. $\mathbf{0} \leq \mathbf{x} \leq \mathbf{1} \equiv 0 \leq x_i \leq 1$, $i=1,2,...n$,), where vectors $\mathbf{0}$:=(0,0,...,0),

1:=(1,1,...,1) will be called model vectors. This follows from the fact that variables can be normalized using the operation of zero notarization or a quotient mapping using the maximal value. Values of a set of features assigned to the objects can be then represented on a radar chart. In order to do this let us denote by x_i those points on axis O_i that lie on the intersection of axis O_i and a circle with the centre located at the origin of the coordinate system and a radius of x_i , $i=1,2,...,n$. By joining subsequent points x_1 with x_2, x_2 with x_3, ..., x_n with x_1 we create an n-polygon, which area $S_{\mathbf{x}}$ defined by:

$$S_{\mathbf{X}} = \sum_{i=1}^{n} \frac{1}{2} x_i x_{i+1} \sin\frac{2\pi}{n} = \frac{1}{2}\sin\frac{2\pi}{n}\sum_{i=1}^{n} x_i x_{i+1}, \quad \text{where } x_{n+1} := x_1. \tag{2}$$

The area of an inscribed, regular n-polygon induced by a model vector $\mathbf{1} = (1,1,\ldots,1)$ is defined by

$$S_{\mathbf{1}} = \sum_{i=1}^{n} \frac{1}{2} \cdot 1 \cdot 1 \cdot \sin\frac{2\pi}{n} = \frac{1}{2} n \sin\frac{2\pi}{n},$$

And the ration of those areas $S_{\mathbf{x}}/S_{\mathbf{1}}$ is defined as

$$S_{\mathbf{x/1}} := S_{\mathbf{x}} / S_{\mathbf{1}} = \frac{1}{n}\sum_{i=1}^{n} x_i x_{i+1} \quad \text{where } x_{n+1} := x_1. \tag{3}$$

In further analysis let us assume that for equivalent features we have designated weight coefficients: $w_1, w_2, ..., w_n$, that make up a weight vector:

$$\mathbf{w} := (w_1, w_2, \ldots, w_n) \in \Omega,$$

i.e. $\mathbf{w} > \mathbf{0}, \sum_{i=1}^{n} w_i = 1.$

Let

$$w_{max} = \max_{1 \le i \le n}(w_1, w_2, \ldots, w_n).$$

In a similar manner as above we can create radar charts for vectors $\mathbf{B_w}(\mathbf{x}), \mathbf{B_w}(\mathbf{1})$, calculate the areas of n-polygons that are generated by them: $S_{\mathbf{B_w}(\mathbf{x})}, S_{\mathbf{B_w}(\mathbf{1})}$:

$$S_{B_w(x)} = \sum_{i=1}^{n} \frac{1}{2} \frac{w_i}{w_{\max}} x_i \frac{w_{i+1}}{w_{\max}} x_{i+1} \sin\frac{2\pi}{n} = \frac{1}{2}\sin\frac{2\pi}{n}(w_{\max})^{-2}\sum_{i=1}^{n} w_i w_{i+1} x_i x_{i+1},$$

$$S_{B_w(\mathbf{1})} = \sum_{i=1}^{n} \frac{1}{2} \cdot \frac{w_i}{w_{max}} \cdot 1 \cdot \frac{w_{i+1}}{w_{max}} \cdot 1 \cdot \sin\frac{2\pi}{n} = \frac{1}{2}\sin\frac{2\pi}{n}(w_{max})^{-2}\sum_{i=1}^{n} w_i w_{i+1},$$

where $x_{n+1} := x_1, w_{n+1} := w_1$, operator $\mathbf{B_w}$ is defined by equation (1).From that is follows that the ratio $S_{\mathbf{B_w}(\mathbf{x})} / S_{\mathbf{B_w}(\mathbf{1})}$ of areas of above polygons is defined by:

$$S^{\mathbf{w}}_{\mathbf{x}/\mathbf{1}} := S_{\mathbf{B_w}(\mathbf{x})} / S_{\mathbf{B_w}(\mathbf{1})} = \frac{\sum_{i=1}^{n} w_i w_{i+1} x_i x_{i+1}}{\sum_{i=1}^{n} w_i w_{i+1}} \qquad (4)$$

Let us consider an object **Q** which features allow it to be described by a vector **x**= $(x_1,x_2,...,x_n) \in \Re^n_+$. If **x** has only one coordinate not equal to 0, that is $p\in$N: $1\le p\le n$, $x_p\neq 0$ and $x_i=0$ for $i\neq p$ then we shall call this vector a **singular** vector. If **x** is a singular vector then vector $\mathbf{B_w x}$, as defined by (1) is also a singular vector. A radar chard induced by a singular vector is only line segment with an area of 0. Considering the above, for definitiveness, let us assume that analyzed vectors are not singular. This can be achieved in many ways, for example by changing the coordinate system, of which a normalization of variables induced by those vectors is a special case. Another example of a change of a coordinate system is an affine transformation of the space $\Re^n$ into itself by [Stark 1970]

$$\mathbf{x}' = \mathbf{Ax} + \mathbf{b},$$

where the matrix:

$\mathbf{A} = \left[a_{ij}\right]_{n\times n}$, determinant $|\mathbf{A}| \neq 0$, vector $\mathbf{b}^T = \left[b_1, b_2, ..., b_n\right]$.

This approach allows for an intuitive visualization of individual objects, considering the importance of individual components (values of features). Let us illustrate this by a simple example posted in Table 1.

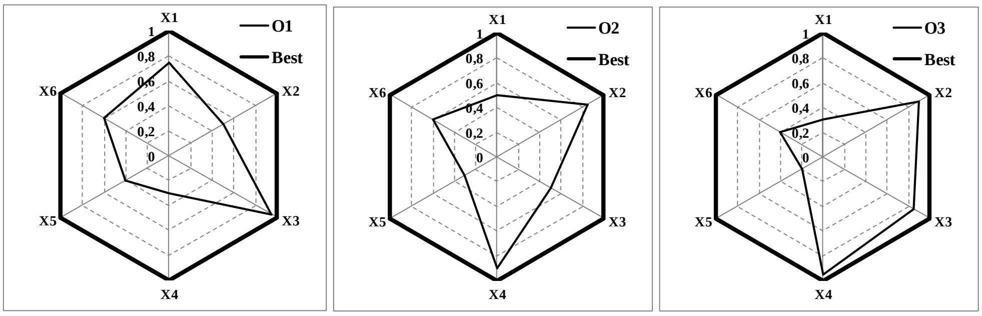
Table 1. The example's values of six variables of four objects in their original form (left hand side) and values of variables after they have been transformed by a weight operator defined by (1), induced by weight **w**=(0.225, 0.1, 0.1, 0.1, 0.25, 0.225) (right hand side).

	X1	X2	X3	X4	X5	X6		X1	X2	X3	X4	X5	X6
O1	0,75	0,5	0,95	0,3	0,4	0,6	**O1***	0,68	0,2	0,38	0,12	0,4	0,5
O2	0,5	0,85	0,5	0,9	0,3	0,6	**O2***	0,45	0,34	0,2	0,36	0,3	0,5
O3	0,3	0,9	0,85	0,95	0,2	0,4	**O3***	0,27	0,36	0,34	0,38	0,2	0,4
Best	1	1	1	1	1	1	**Best***	0,9	0,4	0,4	0,4	1	0,9

Source: own research

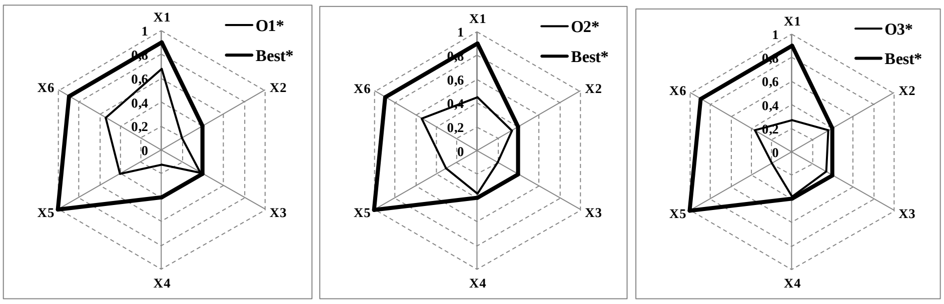
EXAMPLE 1.In the below pictures – Fig. 1, Fig. 2, we have presented visualizations of three objects defined by Table 1., with the forth object **Best** serving as the background. Object **Best** is evaluated as the best object considering each and every of its features. Those features can be interpreted as stimulators. It can be clearly seen that areas of objects **O1**, **O2** and **O3** in Figure 1 create an increasing sequence. However, when we consider the values of individual variables expressed by the weight vector **w**=(0.225, 0.1, 0.1, 0.1, 0.25, 0.225), then as shown in Figure 2, areas of objects **O1***, **O2*** and **O3*** creates a decreasing sequence. Moreover, let us note that while evaluating the analyzed phenomenon the radar chart of the dominant object **Best*** in Figure 2 clearly shows which variables are most important. If we were to use equations (3) and (4) to evaluate objects **O1**, **O2** and **O3**, then in a case when we consider all variables uniformly important we would obtain results of 0.32, 0.34, 0.37, respectively, for (3), and 0.32, 0.29, 0.21, respectively, for (4), when the evaluation was done using the weight vector **w**.

Figure 1. Radar charts for objects defined in Table 1 (left hand side)



Source: own research

Figure 2. Radar charts for objects defined in Table 1 (right hand side)



Source: own research

Let us note the formulas (3) and (4) use the intuition connected with visually representing objects as radar charts and even for a large number of features have the defect of depending on the ordering of those features. Which means that they are sensitive to changes in order of considered features which in many situation may result in a different ordering of evaluated objects.

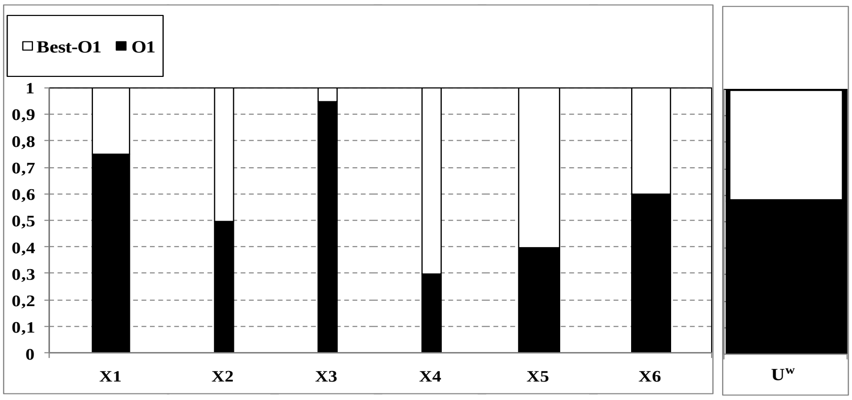
For example, if we were to order the variables form example 1 as: X1, X2, X5, X3,X6, X4, then both the graphical representation in Pic. 2, and the values obtained from (4) would result in a different ordering of objects. Indeed, the values obtained from (4) for objects **O1**, **O2** and **O3** would be 0.32, 0.35, 0.27, respectively, as opposed to previous values of 0.32, 0.29, 0.21.

In those cases when there are no instrumental limitations when creating reports one can imagine different ways of visualizing objects defined as sets of features. One can use, for example, an illustration such as one in Pic. 3. For the object **O1** from Table 1 and the vector of weights **w,** the illustration shows the levels of liquids in test tubes, as the information about both the importance of individual features and the value of the popular synthetic indicator **U,** defined by the formula [Cieślak 1974, Kukuła 2000]:

$$\mathbf{U}^{\mathrm{w}} = \sum_{i=1}^{n} w_i x_i \,, \tag{5}$$

where $\mathbf{w} = (w_1, w_2, ..., w_n), \mathbf{x} = (x_1, x_2, ..., x_n) \in \Re_+^n = \{\mathbf{y} : \mathbf{y} \geq 0\}, \sum_{i=1}^{n} w_i = 1.$

Figure 3. Graphical illustration of object **O1** from Table 1with object **Best** in the background



Source: own research

Unfortunately, this kind of visualization is not supported out of the box by the tooling available within a spreadsheet. Another way of representing data, especially useful when dealing with large numbers of variables and objects are overrepresentation maps [see Szczesny et al. 2012, Binderman, Borkowski, Szczesny, Shachmurove 2012]. In the following section we shall demonstrate how to obtain radar coefficients that are independent from the ordering of variables.

SYNTHETIC RADAR MEASURES

In order to construct radar weight measures that are independent from the ordering of features, let us consider an object **Q**, which features allow to describe it by a vector $\mathbf{x}=(x_1,x_2,...,x_n) \in \Re_+^{\mathrm{n}}$ and a weight $\mathbf{w} := (w_1, w_2, \ldots, w_n) \in \Omega$.

Let us denote the j-th permutation of the set of coordinates of the vector $B_{\mathbf{w}}\mathbf{x} := (w_{max})^{-1}(w_1x_1, w_2x_2, \dots, w_nx_n)$.

By

$\mathbf{x}^{\mathbf{w}}{}_j := (x^w{}_{1j}, x^w{}_{2j}, \dots, x^w{}_{nj}) = (w_{max})^{-1}\left(w_{j1}x_{j1}, w_{j2}x_{j2}, \dots, w_{jn}x_{jn}\right)$, $(\mathbf{x}^{\mathbf{w}}{}_1 := B_{\mathbf{w}}\mathbf{x})$,

where $j = 1,2,\dots,n!$.

Let us assume the following (see (4)):

$$S^{\mathbf{w}}(\mathbf{x}_j) := \sqrt{\frac{\sum_{i=1}^{n} w_{ji}w_{ji+1}x_{ji}x_{ji+1}}{\sum_{i=1}^{n} w_{ji}w_{ji+1}}} = \left(\sum_{i=1}^{n} w_{ji}w_{ji+1}\right)^{-\frac{1}{2}} \sqrt{\sum_{i=1}^{n} w_{ji}w_{ji+1}x_{ji}x_{ji+1}} \tag{6}$$

$$\mathrm{M}^{\mathbf{w}}(\mathbf{x}) := \max_{1\le j\le k} \mathrm{S}^{\mathbf{w}}(\mathbf{x}_j), \tag{7}$$

$$\mathrm{S}^{\mathbf{w}}(\mathbf{x}) := \frac{1}{k}\sum_{j=1}^{k} \mathrm{S}^{\mathbf{w}}(\mathbf{x}_j), \tag{8}$$

$$\mathrm{m}^{\mathbf{w}}(\mathrm{x}) := \min_{1\le j\le k} \mathrm{S}^{\mathbf{w}}(\mathbf{x}_j), \tag{9}$$

where $k:=n!$.

NOTE 1. If $\mathbf{x}=(x_1,x_2,\dots,x_n) \in \Re^n_+$ is a singular vector and j-th variable $x_j \neq 0$ then it's synthetic natural measure should be defined as $\frac{w_jx_j}{n}$.

NOTE 2. The defined above measures $M^w(\mathbf{x})$, $S^w(\mathbf{x})$ and $m^w(\mathbf{x})$ are **independent** from the ordering of coordinates of vector **x.** In case of object **O1** which was considered in Example 1, we have

$S^w(\mathbf{x})=0{,}56$; $M^w(\mathbf{x}) = 0{,}62$; $\mathrm{S}^w(\mathbf{x})=0{,}57$; $m^w(\mathbf{x})=0{,}53$.

The methods of calculating synthetic measures that were defined above fulfil the basic postulate of stability of the used method [see Jackson 1970] – they are independent from the ordering of features that describe the given object. Radar methods may seem as computationally complex, but in the computer age this is no longer an issue, as they can be easily calculated in a spreadsheet – naturally, for a limited set of variables.

Let $\mathbf{w} := (w_1, w_2, \dots, w_n) \in \Omega$, $\mathbf{x} \in \Re^n_+, \alpha \in \Re_+$. Functions that were defined above have the following properties (with consideration of Note 1) [see Binderman, Borkowski, Szczesny 2008, Binderman, Szczesny 2009]:

a) $M^w(\mathbf{0})=S^w(\mathbf{0})=m^w(\mathbf{0})=0$,

b) $M^w(\mathbf{1})=S^w(\mathbf{1})=m^w(\mathbf{1})=1$,

c) $M^w(\mathbf{1/2})=S^w(\mathbf{1/2})=m^w(\mathbf{1/2})=1/2$,

d) $M^w(\alpha\mathbf{x}) = \alpha M^w(\mathbf{x}),\ S^w(\alpha\mathbf{x}) = \alpha S^w(\mathbf{x}),\ m^w(\alpha\mathbf{x}) = \alpha m^w(\mathbf{x})$,

$m^w(\mathbf{x}) \le S^w(\mathbf{x}) \le M^w(\mathbf{x})$.

The authors have been researching into the subject of measuring similarity or dissimilarity of objects, with special attention given to the area of structural changes in agriculture for multiple years (see [Binderman et al. 2012, Binderman, Borkowski, Szczesny 2008, 2009, 2010a,b,c, 2011, 2012a Binderman, Szczesny 2009, 2011, Borkowski, Szczesny 2002]). Literature on this subject provides a wide array of instruments to compare and contrast objects represented by vectors. The authors in their other works use mainly measures based on visual representations via radar charts.

RESULTS OF EMPIRICAL RESEARCH

In the analysis of regional dissimilarity of agriculture in Poland, synthetic measures prove to be a fundamental tool. For they allow to describe analyzed objects (voivodeships) characterized by vectors from a multidimensional space of features by means of a single synthetic measure. By using this method one can perform comparative analysis of voivodeships with special consideration given to ordering of objects according to a predefine criterion and defining the place at which a given voivodeship ranks in comparison to other analyzed objects. Moreover, when analyzing a given time period one can define both the direction of change as well as the change in the level of dissimilarity in a group of objects. To illustrate the ideas and mechanisms described in previous sections the authors have used empirical data of Polish agriculture in an area formal (divided into voivodeships) for the period of 2001 – 2010. In order to eliminate or at least limit small seasonal deviations the period has been divided into three sub periods: 2001 – 2003, 2005 – 2007 and 2008 – 2010. Afterwards the data for those sub periods was averaged using CPI coefficient to realize the values given in PLN. 2004 has been eliminated from the analysis as in this year Poland joined the UE.
The following six features were used in the analysis:

1. GNP (X1) – Gross National Product (current prices) per capita given in PLN,
2. APG (X2) – Agricultural Production of Goods (PLN/ha),
3. AAH (X3) – Average Area of a single agricultural Household in hectares,
4. YIELD (X4) – yield of four crops in tones per hectare
5. UMF (X5) – Usage of Mineral Fertilizers in kilogram's per hectare
6. EMP (X6) – the number of people employed in agriculture, hunting and forestry, fishing for 100 hectares of general area.

All the features apart from X6 have had the character of a stimulator, X6 was assumed as a de-stimulator. The chosen features do not exhaust all aspects

of agriculture, but on a general level, allow to show the differences which appear between voivodeships.
Based on the values of diagnostic variables in individual sub periods, we have created two constant (static) hypothetical voivodeships: minimal $\mathbf{Q}_0$ and maximal $\mathbf{Q}_{49}$, described by the least and most advantageous set of feature values, respectively. Those voivodeships, as models, will be denoted by vectors $\mathbf{x}_0$ and $\mathbf{x}_{49}$, each having 6 dimensions.

Thus we have 50 objects $\mathbf{Q}_0$, $\mathbf{Q}_1$, $\mathbf{Q}_2$,..., $\mathbf{Q}_{49}$, each described by n = 6 features $X_1, X_2, ..., X_6$. As the chosen diagnostic variables had different units and orders of magnitude, they were all normalized. In order to make the variables comparable, from a set of approaches, the method of zero notarization was chosen [Kukula, 2000], converting the variables according to the following:

$$z_{ij} = \frac{x_{ij} - x_{mj}}{x_{Mj} - x_{mj}} \quad - \text{for stimulators, } j = 1,2,3,4,5;$$

$$z_{ij} = \frac{x_{Mj} - x_{ij}}{x_{Mj} - x_{mj}} \quad - \text{for de-stimulators, } j = 6, \qquad (10)$$

where $x_{mj} = \min\limits_{1 \le i \le 48} x_{ij}$, $x_{Mj} = \max\limits_{1 \le i \le 48} x_{ij}$, $0 \le i \le 49,\ 1 \le j \le 6;$

The applied transformation can be represented symbolically as:

$$\mathbf{Z} = \varphi(\mathbf{X}),$$

where $\mathbf{X}$ – observation matrix, $\mathbf{Z} = [z_{ij}]_{48\times 6}$.

After the transformation of variables, the static model vectors are as follows:

$$\mathbf{z}_0 = \mathbf{0} = [0,0,...,0], \quad \mathbf{z}_{49} = \mathbf{1} = [1,1,...,1].$$

Additionally, one should note that it is non-trivial to show that they are equally important. Especially, GNP describes all economic activities, not just agricultural one, which input into GNP is at a level of a few percent. Yields of crops are largely dependent on the quality of soil and general environment conditions. Thus, is a comparative analysis, this fact should be noted by assigning different weights to features. So, to illustrate the problem, for the purposes of this work the following set of weights was assumed:

$$\mathbf{w} = (0{,}05;\ 0{,}225;\ 0{,}275;\ 0{,}05;\ 0{,}05;\ 0.35). \qquad (11)$$

Empirical data, after normalization is shown in the following Table 1.

Table 1. Input data after normalization via zero unitarization (Formula 10)

	2001-2003						2005-2007						2008-2010					
	X1	X2	X3	X4	X5	X6	X1	X2	X3	X4	X5	X6	X1	X2	X3	X4	X5	X6
DOLNOŚLĄSKIE	0,20	0,16	0,45	0,58	0,23	0,87	0,38	0,21	0,44	0,55	0,57	0,89	0,54	0,29	0,49	0,72	0,68	0,88
KUJAWSKO-POMORSKIE	0,13	0,45	0,59	0,33	0,54	0,70	0,23	0,49	0,61	0,32	0,88	0,76	0,31	0,58	0,67	0,44	0,78	0,80
LUBELSKIE	0,00	0,16	0,22	0,15	0,25	0,44	0,08	0,24	0,22	0,07	0,45	0,51	0,16	0,37	0,24	0,27	0,35	0,59
LUBUSKIE	0,11	0,02	0,48	0,10	0,33	0,91	0,25	0,12	0,55	0,13	0,50	0,86	0,31	0,25	0,66	0,32	0,38	0,92
ŁÓDZKIE	0,13	0,49	0,23	0,03	0,26	0,58	0,27	0,74	0,24	0,08	0,62	0,61	0,38	0,75	0,25	0,21	0,48	0,64
MAŁOPOLSKIE	0,09	0,22	0,00	0,18	0,15	0,00	0,22	0,39	0,00	0,23	0,12	0,16	0,32	0,37	0,02	0,29	0,04	0,39
MAZOWIECKIE	0,55	0,38	0,27	0,05	0,14	0,59	0,81	0,61	0,28	0,00	0,39	0,75	1,00	0,73	0,30	0,12	0,28	0,74
OPOLSKIE	0,06	0,33	0,45	0,73	0,59	0,77	0,19	0,40	0,46	0,76	0,82	0,79	0,28	0,59	0,53	1,00	1,00	0,86
PODKARPACKIE	0,00	0,00	0,02	0,17	0,02	0,34	0,08	0,03	0,03	0,15	0,04	0,41	0,16	0,00	0,04	0,28	0,00	0,40
PODLASKIE	0,04	0,16	0,53	0,01	0,17	0,67	0,13	0,32	0,55	0,00	0,24	0,78	0,20	0,41	0,57	0,11	0,29	0,81
POMORSKIE	0,19	0,08	0,59	0,25	0,52	0,88	0,32	0,25	0,64	0,29	0,52	0,87	0,41	0,45	0,69	0,39	0,52	0,91
ŚLĄSKIE	0,24	0,28	0,06	0,34	0,25	0,77	0,39	0,43	0,09	0,28	0,40	0,75	0,52	0,54	0,11	0,39	0,36	0,83
ŚWIĘTOKRZYSKIE	0,04	0,27	0,13	0,04	0,10	0,44	0,14	0,40	0,11	0,02	0,35	0,26	0,25	0,45	0,12	0,15	0,19	0,44
WARMIŃSKO-MAZURSKIE	0,04	0,18	0,87	0,23	0,18	0,93	0,14	0,27	0,93	0,18	0,48	0,92	0,21	0,36	1,00	0,47	0,41	0,92
WIELKOPOLSKIE	0,22	0,64	0,53	0,31	0,36	0,68	0,38	0,86	0,51	0,32	0,76	0,79	0,49	1,00	0,53	0,43	0,72	0,78
ZACHODNIOPOMORSKIE	0,16	0,04	0,71	0,26	0,39	0,98	0,26	0,13	0,94	0,30	0,45	0,98	0,34	0,23	1,00	0,54	0,47	1,00

Source: own calculation based on data published by GUS

Based on the above data, the synthetic measures were calculated using the weight vector as defined in (11) and their results placed in Table 2. Table 3 shows the ranking of voivodeships according to those measures. Moreover, a division into four groups was performed according to E. Nowak's method [Nowak 1990], and its results places in Table 4.

Table 2. Values of synthetic coefficients as defined by (5) – (9), calculated using the weight vector as defined by (11). U* denotes the popular coefficient defined by (5) with the assumption that all weight are identical

	2001-2003						2005-2007						2008-2010					
	U*	U	S	$\mathcal{M}$	$\mathcal{S}$	m	U*	U	S	$\mathcal{M}$	$\mathcal{S}$	m	U*	U	S	$\mathcal{M}$	$\mathcal{S}$	m
DOLNOŚLĄSKIE	0,42	0,52	0,35	0,57	0,44	0,33	0,51	0,55	0,45	0,58	0,50	0,42	0,60	0,61	0,54	0,64	0,58	0,50
KUJAWSKO-POMORSKIE	0,46	0,56	0,48	0,59	0,53	0,43	0,55	0,62	0,55	0,66	0,59	0,52	0,60	0,67	0,61	0,70	0,65	0,58
LUBELSKIE	0,20	0,27	0,19	0,28	0,24	0,18	0,26	0,32	0,26	0,33	0,29	0,22	0,33	0,40	0,32	0,42	0,37	0,31
LUBUSKIE	0,32	0,48	0,26	0,56	0,35	0,17	0,40	0,52	0,37	0,61	0,44	0,31	0,47	0,61	0,45	0,67	0,54	0,45
ŁÓDZKIE	0,29	0,40	0,31	0,44	0,35	0,23	0,43	0,49	0,43	0,58	0,46	0,36	0,45	0,52	0,45	0,60	0,49	0,40
MAŁOPOLSKIE	0,11	0,07	0,05	0,10	0,06	0,02	0,19	0,18	0,13	0,22	0,16	0,09	0,24	0,25	0,18	0,33	0,22	0,12
MAZOWIECKIE	0,33	0,40	0,36	0,44	0,37	0,29	0,47	0,54	0,51	0,62	0,50	0,40	0,53	0,58	0,56	0,67	0,55	0,44
OPOLSKIE	0,49	0,54	0,44	0,57	0,50	0,42	0,57	0,58	0,51	0,62	0,56	0,50	0,71	0,69	0,64	0,75	0,68	0,63
PODKARPACKIE	0,09	0,13	0,03	0,12	0,06	0,01	0,12	0,17	0,09	0,15	0,11	0,08	0,15	0,17	0,10	0,19	0,12	0,04
PODLASKIE	0,26	0,43	0,25	0,49	0,34	0,16	0,34	0,52	0,36	0,56	0,44	0,24	0,40	0,56	0,43	0,60	0,50	0,34
POMORSKIE	0,42	0,54	0,36	0,63	0,45	0,29	0,48	0,59	0,46	0,66	0,53	0,44	0,56	0,68	0,57	0,72	0,63	0,55
ŚLĄSKIE	0,32	0,39	0,27	0,41	0,31	0,23	0,39	0,44	0,35	0,50	0,38	0,28	0,46	0,50	0,40	0,58	0,45	0,33
ŚWIĘTOKRZYSKIE	0,17	0,26	0,17	0,28	0,22	0,12	0,21	0,24	0,21	0,28	0,22	0,17	0,27	0,32	0,26	0,37	0,29	0,23
WARMIŃSKO-MAZURSKIE	0,40	0,63	0,36	0,74	0,50	0,30	0,48	0,68	0,48	0,79	0,58	0,42	0,56	0,73	0,56	0,83	0,66	0,52
WIELKOPOLSKIE	0,46	0,57	0,51	0,60	0,54	0,42	0,60	0,68	0,63	0,74	0,66	0,58	0,66	0,73	0,69	0,79	0,71	0,64
ZACHODNIOPOMORSKIE	0,42	0,59	0,34	0,71	0,46	0,25	0,51	0,68	0,44	0,82	0,57	0,39	0,60	0,74	0,55	0,87	0,66	0,51
min	0,091	0,069	0,034	0,096	0,056	0,008	0,124	0,171	0,092	0,146	0,108	0,085	0,147	0,175	0,102	0,195	0,116	0,040
max	0,488	0,626	0,506	0,741	0,542	0,426	0,604	0,684	0,631	0,816	0,661	0,576	0,711	0,744	0,690	0,867	0,709	0,639
max-min	0,397	0,557	0,472	0,644	0,487	0,418	0,480	0,513	0,539	0,670	0,553	0,491	0,564	0,569	0,588	0,673	0,593	0,598
mi	0,322	0,423	0,295	0,470	0,358	0,239	0,407	0,487	0,390	0,544	0,438	0,337	0,474	0,548	0,457	0,608	0,505	0,413
sigma	0,123	0,160	0,131	0,187	0,147	0,125	0,140	0,167	0,146	0,192	0,159	0,143	0,156	0,172	0,162	0,184	0,171	0,169
V	0,381	0,378	0,445	0,398	0,410	0,521	0,344	0,343	0,374	0,354	0,362	0,424	0,330	0,313	0,356	0,303	0,338	0,410

Source: own research

In the above Table 2 mi, sigma, V denote average value, standard deviation, volatility, respectively. Values in columns U^w, S^w, M^w, S^w, m^w are values of synthetic measures calculated according to (5) – (9), using the weight vector as defined by (11), while the values in column U* are the values of a synthetic measures as defined by (5) but with the assumption that all weights are identical, that is the weight vector !! $\mathbf{w} = \mathbf{1/n} = (1/n, 1/n, ..., 1/n)$!!.

Table 3. Division of voivodeships into four groups according to the values of coefficients from Table 2

	2001-2003						2005-2007						2008-2010					
	U*	U	S	$\mathcal{M}$	$\mathcal{S}$	*m*	U*	U	S	$\mathcal{M}$	$\mathcal{S}$	*m*	U*	U	S	$\mathcal{M}$	$\mathcal{S}$	*m*
DOLNOŚLĄSKIE	6	7	7	7	7	4	5	7	7	9	8	6	3	8	8	9	7	7
KUJAWSKO-POMORSKIE	2	4	2	5	2	1	3	4	2	5	2	2	4	6	3	6	5	3
LUBELSKIE	13	13	13	13	13	11	13	13	13	13	13	13	13	13	13	13	13	13
LUBUSKIE	9	8	11	8	10	12	10	9	10	8	11	10	9	7	9	8	9	8
ŁÓDZKIE	11	11	9	10	9	9	9	11	9	10	9	9	11	11	10	10	11	10
MAŁOPOLSKIE	15	16	15	16	16	15	15	15	15	15	15	15	15	15	15	15	15	15
MAZOWIECKIE	8	10	6	11	8	6	8	8	4	7	7	7	8	9	6	7	8	9
OPOLSKIE	1	5	3	6	3	3	2	6	3	6	5	3	1	4	2	4	2	2
PODKARPACKIE	16	15	16	15	15	16	16	16	16	16	16	16	16	16	16	16	16	16
PODLASKIE	12	9	12	9	11	13	12	10	11	11	10	12	12	10	11	11	10	11
POMORSKIE	5	6	4	3	6	7	7	5	6	4	6	4	6	5	4	5	6	4
ŚLĄSKIE	10	12	10	12	12	10	11	12	12	12	12	11	10	12	12	12	12	12
ŚWIĘTOKRZYSKIE	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14
WARMIŃSKO-MAZURSKIE	7	1	5	1	4	5	6	3	5	2	3	5	7	2	5	2	4	5
WIELKOPOLSKIE	3	3	1	4	1	2	1	1	1	3	1	1	2	3	1	3	1	1
ZACHODNIOPOMORSKIE	4	2	8	2	5	8	4	2	8	1	4	8	5	1	7	1	3	6

Source: own research

Table 4. Ranking of voivodeships according to coefficients from Table 2

	2001-2003						2005-2007						2008-2010					
	U*	U	S	$\mathcal{M}$	$\mathcal{S}$	*m*	U*	U	S	$\mathcal{M}$	$\mathcal{S}$	*m*	U*	U	S	$\mathcal{M}$	$\mathcal{S}$	*m*
DOLNOŚLĄSKIE	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
KUJAWSKO-POMORSKIE	1	2	1	2	1	1	1	2	1	2	2	1	2	2	2	2	2	2
LUBELSKIE	3	3	3	3	3	3	4	3	3	4	3	3	3	3	3	4	3	3
LUBUSKIE	2	2	3	2	3	3	3	2	3	2	2	3	3	2	3	2	2	2
ŁÓDZKIE	3	3	2	3	3	3	2	2	2	2	2	2	3	3	3	3	3	3
MAŁOPOLSKIE	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
MAZOWIECKIE	2	3	2	3	2	2	2	2	2	2	2	2	2	2	2	2	2	2
OPOLSKIE	1	2	1	2	2	1	1	2	2	2	2	1	1	2	1	2	1	1
PODKARPACKIE	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
PODLASKIE	3	2	3	2	3	3	3	2	3	2	2	3	3	2	3	3	3	3
POMORSKIE	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
ŚLĄSKIE	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
ŚWIĘTOKRZYSKIE	4	4	3	4	3	3	4	4	4	4	4	4	4	4	4	4	4	4
WARMIŃSKO-MAZURSKIE	2	1	2	1	2	2	2	1	2	1	2	2	2	1	2	1	2	2
WIELKOPOLSKIE	1	2	1	2	1	1	1	1	1	1	1	1	1	1	1	2	1	1
ZACHODNIOPOMORSKIE	2	1	2	1	2	2	2	1	2	1	2	2	2	1	2	1	2	2

Source: own research

Analysis of results in Tables 2 – 4 has shown that the level of agriculture is spatially variable and the division of voivodeships into four groups according to analyzed features is in accordance with regional dispersion. The chosen example is a set of sixteen objects, which differ significantly between each other. In those cases, every method will result in a similar ordering of objects. That is why all coefficients indicate that the average level increases between subsequent periods. However, this cannot be stated when differentiation is considered (standard deviation has a different trend than volatility). The results show that even in this relatively stable example the measure of ordering S^w, defined by (6) return with one of the worst (random) orderings. Those are significant in cases of a few voivodeships and the magnitude of inconsistency between the ordering and other coefficients is comparable only to that of inconsistency represented by coefficient U*. However, that coefficient is using a completely different set of weights (so it is based on a different valuation criteria). Research has shown that the best ordering is given by a method based on coefficient defined by (7). The results are relatively close to those returned when using coefficient U, but, as mentioned before, when objects are so different between each other, most methods should lead to similar orderings.

SUMMARY

In reports generated by information system, graphical representations are desirable as they depict in an easy to understand and process way even very difficult ideas by using the recipients intuition and not requiring them to thoroughly prepare in terms of both knowledge and toolkit. However, improper use can lead to misunderstandings and incorrect conclusions and, as a result, decline of trust in the information system. In this work it has been shown that preparing radar charts requires special care. Charts that are a promising and easy-to-access tool available in a spreadsheet, which allow to present and evaluate objects described by many features. Analysis performed on data describing the level of agriculture between different voivodeships has shown that used methods of classification and ordering of objects with six features, even though the objects were very dissimilar, have given different results. Moreover, it has been noted that the very intuitive synthetic coefficient S^w that uses the area of an n-polygon is not constant in relation to the ordering of variables, which can, in many cases, be a large disadvantage. A similar behavior can be observed when considering the weighting process. Coefficient U* which considers all variables equally and coefficient U^w which takes weights into account returned with different results, even though the base material was very diverse.

Because of this it seems necessary to place in visualization reports results of individual objects (processes) in a multidimensional view (for example by using radar charts with an ideal and “desired” or “expected” objects in the background) of multiple values, rather than a single synthetic coefficient which evaluates the given object as the most significant one. In case when they will return with different evaluation the recipient will, at the very least, investigate the causes of the discrepancies. It is especially important when the report consists of a set of hierarchical visualizations and evaluations, because then unexpected situations on lower levels can go unnoticed.

REFERENCES

Binderman, Z., Borkowski B., Szczesny W. (2012a) Radar coefficient of concentration, Quantitative Methods in Economics Vol. XIII, no 2, pp. 7-21.

Binderman, Z., Borkowski B., Szczesny W. Shachmurove Y. (2012) Zmiany struktury eksportu produktów rolnych w wybranych krajach UE w okresie 1980-2010, Quantitative Methods in Economics Vol. XIII, nr 1, pp. 36-48.

Binderman, Z., Borkowski B., Szczesny W. (2011) An application of radar charts to geometrical measures of structures' of conformability, Quantitative Methods in Economics Vol. XII, nr 1, pp. 1-13.

Binderman Z., Borkowski B., Szczesny W. (2010a) Radar measures of structures' conformability, Quantitative methods in economy XI, nr 1, 1-14.

Binderman Z., Borkowski B., Szczesny W. (2010b) Analiza zmian struktury spożycia w Polsce w porównaniu z krajami unii europejskiej. Metody wizualizacji danych w analizie zmian poziomu i profilu konsumpcji w krajach UE, , RNR PAN, Seria G, Ekonomika Rolnictwa , T. 97, z. 2, pp. 77-90.

Binderman Z., Borkowski B., Szczesny W. (2010c) The tendencies in regional differentiation changes of agricultural production structure in Poland, Quantitative methods in regional and sectored analysis, U.S., Szczecin, pp. 67-103.

Binderman Z., Borkowski B., Szczesny W. (2009) Tendencies in changes of regional differentiation of farms structure and area, Quantitative methods in regional and sectored analysis/sc., U.S., Szczecin, pp. 33-50.

Binderman Z., Borkowski B., Szczesny W. (2008) O pewnej metodzie porządkowania obiektów na przykładzie regionalnego zróżnicowania rolnictwa, Metody ilościowe w badaniach ekonomicznych, IX, pp. 39-48, wyd. SGGW.

Binderman Z., Szczesny W. (2009) Arrange methods of tradesmen of software with a help of graphic representations Computer algebra systems in teaching and research, Siedlce Wyd. WSFiZ, pp. 117-131.

Binderman Z., Szczesny W. (2011) Comparative Analysis of Computer Techniques for Visualization Multidimensional Data, Computer algebra systems in teaching and research, Siedlce, wyd. Collegium Mazovia, pp. 243-254.

Binderman Z. (2011) Matematyczne aspekty metod radarowych, Metody ilościowe w badaniach ekonomicznych, XII, nr 2, pp. 69-79.

Borkowski B., Szczesny W. (2005) Metody wizualizacji danych wielowymiarowych jako narzędzie syntezy informacji, SERiA, Roczniki Naukowe, t. VII, pp. 11-16.

Borys T. (1984) Kategorie jakości w statystycznej analizie porównawczej, AE Wrocław, nr. 284, seria MiO, nr 23.

Cieślak M., (1974) Modele zapotrzebowania na kadry kwalifikowane, PWN, Warszawa.

Grabiński T., Wydymus S., Zeliaś A. (1989)Metody taksonomii numerycznej w modelowaniu zjawisk społeczno-gospodarczych, PWN, Warszawa.

Hellwig Z., (1976) Schemat budowy prognozy statystycznej metodą wag harmonicznych, Przegląd Statystyczny, z. 1.

Kukuła K. (2000) Metoda unitaryzacji zerowanej, PWN, Warszawa.

Malina A. (2004) Wielowymiarowa analiza przestrzennego zróżnicowania struktury gospodarki Polski według województw, AE, Seria Monografie nr 162, Kraków.

Młodak A. (2006) Analiza taksonomiczna w statystyce regionalnej, Warszawa.

Stark M. (1970) Geometria Analityczna, Monografie matematyczne 17, PWN, Warszawa.

Nowak E. (1990) Metody taksonomiczne w klasyfikacji obiektów społeczno-gospodarczych, PWE, Warszawa.

Pociecha J., Podolec B., Sokołowski A., Zając K. (1988) Metody taksonomiczne w badaniach społeczno-ekonomicznych, PWN, Warszawa.

Odrzygóźdź, Z., Karwański, M., Szczesny W. (2009) Business Intelligence- Application and trends in development, Information Systems in Management III, WULS Press, Warsaw 2009, 3, pp. 78-87.

Rusnak Z., Siedlecka U., Siedlecki J. (1982) Metody badania ważności cech, AE Wrocław nr 215.

Stobiecka J. (1998) Ustalanie współczynników ważności cech produktów rynkowych, Wiadomości Statystyczne nr 8.

Szczesny W. i inni, (2012) Models and Methods of Grade Data Analysis Recent Developments, Institute of Computer Science Polish Academy of Sciences.

Vesset D., McDonough B., Wardley M., Schubmehl D. (2012) Worldwide Business Analytics Software 2012–2016 Forecast and 2011 Vendor Shares, IDC # 235494,Volume1, TabVendors Business Analytics Solutions Market Analysis.

Zeliaś A. (2000) Taksonomiczna analiza przestrzennego zróżnicowania poziomu życia w Polsce w ujęciu dynamicznym, Kraków.