

SIMILARITIES AND DIFFERENCES IN THE CORE MARKETS FOR AGRICULTURAL PRODUCTS IN POLAND AFTER THE ACCESSION TO THE EUROPEAN UNION*

Andrzej Czyżewski¹✉, Dariusz Czakowski²

¹Poznań University of Economics and Business, Poland

²Kujawy and Pomorze University in Bydgoszcz, Poland

Abstract. This paper attempts to determine the similarity between individual markets for agricultural products, considering the evolution of basic balance sheet data (production, domestic consumption, imports and exports) and real purchase prices of agricultural products. The main markets for agricultural products include those demonstrating the highest production volumes of commodity such as cereals, potatoes, sugar beets (sugar), rape, fruit, vegetables, pork, beef and poultry, cow's milk and eggs in the study period. Based on the search for data sources, and considering the specifics of empirical data collected, it was decided that the following methods will be used to pursue the objectives of this study: principal component analysis, cluster analysis based on the Ward's method and k-means. The research allowed to identify three groups of Polish markets which followed a similar development trend after Poland's accession to the European Union. The first group consisted of markets for cereals, rape, poultry and fruits. The second cluster included markets for cow's milk, eggs and beef. The last one was composed of potato, sugar beet, vegetable and pork markets.

Keywords: European integration, markets for agricultural products

INTRODUCTION

Searching for similarities and grouping repetitive patterns in the surrounding world is an inherent part of human nature. While enabling a gradual reduction of entropy, the simplification of complex processes and phenomena allows to discover some unknown relationships (Marek, 1989). As regards economic sciences, methods for object grouping and sharing are an essential element of many studies and analyses. Merging or dividing objects into clusters is a way to obtain homogeneous groups which, while being similar to each other, differ from external objects by a specific feature (or multiple features). Therefore, this paper attempts to determine the similarity between individual markets for agricultural products, considering the evolution of basic balance sheet data (production, domestic consumption, imports and exports) and real purchase prices of agricultural products. Because of their aggregate nature and comparability, the above variables were found to be best suited to characterize the resources of specific markets.

*Publication costs were financed from Kujawy and Pomorze University in Bydgoszcz.

✉ prof. dr hab. Andrzej Czyżewski, Department of Macroeconomics and Agricultural Economics, Poznań University of Economics and Business, Al. Niepodległości 10, 61-875 Poznań, Poland, e-mail: kmigz@ue.poznan.pl

METHODS

The key aspect in this study was to select an appropriate method and way of grouping compound objects because many statistical methods fail to answer the question on the number of resulting groups (Fraley and Raftery, 1998). Hence, the principal component method (Jolliffe, 2002) and the cluster analysis (Cormack, 1971; MacQueen, 1967), including the Ward's method and k-means, are among the most frequently used multidimensional statistical methods for object grouping. Note also that depending on the purpose of the research, an object may be defined as a single market entity, a sector of the economy or the entire state (Pietrzykowski and Kobus, 2006). The relevant literature provides many examples of the use of object grouping methods with respect to food economy and agricultural markets. This includes the works of the following authors: Błażejczyk-Majka and Kala (2005), Czyżewski (1976), Czyżewski and Czakowski (2017); Czyżewski and Guth (2016), Czyżewski and Strońska-Ziemann (2014), Grzelak (2006), Pietrzykowski and Kobus (2008), Poczta and Pawlak (2008).

The first of the research methods, i.e. principal component analysis, is a technique referred to in statistics as factor analysis. Its main goal is to reduce the number of variables covered by the analysis by introducing synthetic variables (factors) (Krzyszowski, 2000). If the reduction of the number of variables has a relatively small adverse effect on the explanation of total variability by two or three synthetic variables, the differences between objects may be represented graphically (on a two- or three-dimensional grid) (Mądry et al., 2011). On this basis, the principal components method can be used to preliminarily assess the similarity between objects without grouping them, as it does not provide any suggestion on how objects should be grouped (Pietrzykowski and Kobus, 2008). The second method used in this study was the clustering analysis which enables identifying internally homogeneous groups of objects. Depending on the techniques employed as a part of cluster analysis, a distinction is made between agglomerative and deglomerative methods (Marek, 1989). Agglomerative methods, such as the Ward's method, create clusters incrementally by adding subsequent objects to existing groups. In turn, the k-means method (classified as a deglomerative method), requires the target number of groups to be specified in advance. Therefore, it will be used as the last one, and

the number of clusters will be determined based on the results of the Ward's method and the principal components analysis.

The preliminary study relied on data from 1994–2013. This allowed to compare the evolution of the values covered during a ten-year period prior to and following the accession. In the case of methods based on multivariate comparative analyses, one of the main requirements for diagnostic variables is their comparability (Panek and Zwierzchowski, 2013). Therefore, the next step of this research was the standardization of variables. Further during the research procedure, it was decided to take into account the arithmetic mean of standardized variables in the second five-year post-accession period (2009–2013). That time interval was chosen in order to rely on more recent data, and because it provides a clearer picture of the effects of Poland's accession to the European Union than the period immediately following the accession. As an alternative to mean values, the analyses could be based on data from specific years. However, due to significant fluctuations being observed in prices and agricultural production volumes, it was decided to use the mean values which reduce the negative impact of production seasonality and price volatility.

RESULTS OF THE STUDY

The purpose of the research procedure was to identify the similarities and differences in the evolution of selected variables in markets for agricultural products. The initial step was the principal component analysis. The study was based on the arithmetic mean of five standardized primary variables (x_1 : domestic production volume, x_2 : domestic consumption volume, x_3 : export volume, x_4 : import volume, x_5 : real purchase prices) from the 2009–2013 period. The key aspect of the principal component analysis was the selection of the adequate number of factors (principal components). While that decision is made arbitrarily, several techniques are available to support it. The most commonly used criteria are as follows (Panek and Zwierzchowski, 2013): eigenvalues; scree plot; proportion of variance explained; and significance of principal components. In this study, the following were adopted as decisive techniques: the Kaiser's criterion (Kaiser, 1960) which is the basis for retaining the principal components with eigenvalues greater than one; the Cattell criterion (based on

Table 1. Eigenvalues and percentage of the volatility of principal components in the post-accession period

Value number	Eigenvalue	% of total variance	Cumulative eigenvalue	Cumulative %
1	1.768604	35.37208	1.768604	35.3721
2	1.534682	30.69365	3.303286	66.0657
3	0.881099	17.62198	4.184386	83.6877
4	0.533187	10.66375	4.717573	94.3515
5	0.282427	5.64854	5.000000	100.0000

Source: own calculations performed with Statistica ver. 12, the econometric software suite. Based on: GUS, 1994–2013; 1995–2014.

the scree plot analysis); and the cumulative percentage of explained variance of variables covered by the analysis. Therefore, for the next considerations, the first two components were selected, with a cumulative eigenvalue of 3.3, and a cumulative percentage of explanation of the total variation of 66.1% (Table 1). At the same time, it can be concluded that approximately 33.9% of information resources contained in input data were not represented in the further research procedure.

As shown by the correlation matrix between input variables and principal components (Table 2), the first component was the most strongly correlated (among the retained ones) with variables x_1 (domestic production) and x_2 (domestic consumption). The relationship between the synthetic variable and both original variables was negative, which means they had an inhibitory effect. Considering the values represented by the first principal component, it may be defined as ‘domestic resource conditions.’ Therefore, in this study, objects with a greater (lower) value of this component were

characterized by a lower (greater) value of average national production and consumption levels in the post-accession period.

In the last stage of the analysis, the second factor represented all other input variables, i.e.: x_3 (export), x_4 (import) and x_5 (real purchase prices). Considering these variables, the second principal component can be referred to as ‘trade and price conditions.’ Variables x_3 and x_5 demonstrated positive correlation with the second factor while in the case of x_4 the correlation was negative. Therefore, as the value of the second component increases in a given market, so do the trade balance and the level of real purchase prices of agricultural products.

The dispersion of objects in the system of the first two components revealed that each market for agricultural products was somehow similar to another one during the study period (Fig. 1). However, the sugar beet (sugar) and poultry markets were distinct from other objects. The first of them differed because of a very low value of the trade and price factor. This resulted mainly

Table 2. Linear correlation coefficients between original variables and principal components in the post-accession period

Specification	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Domestic production volume (x_1)	-0.690564	0.416000	0.479893	-0.201847	-0.281118
Domestic consumption volume (x_2)	-0.921662	-0.035788	0.062342	0.012747	0.381064
Export volume (x_3)	0.303848	0.785895	-0.222101	-0.466860	0.150858
Import volume (x_4)	0.439074	-0.557823	0.606076	-0.332788	0.134058
Real purchase prices (x_5)	0.396428	0.656923	0.479853	0.404446	0.132137

Source: same as in table 1.

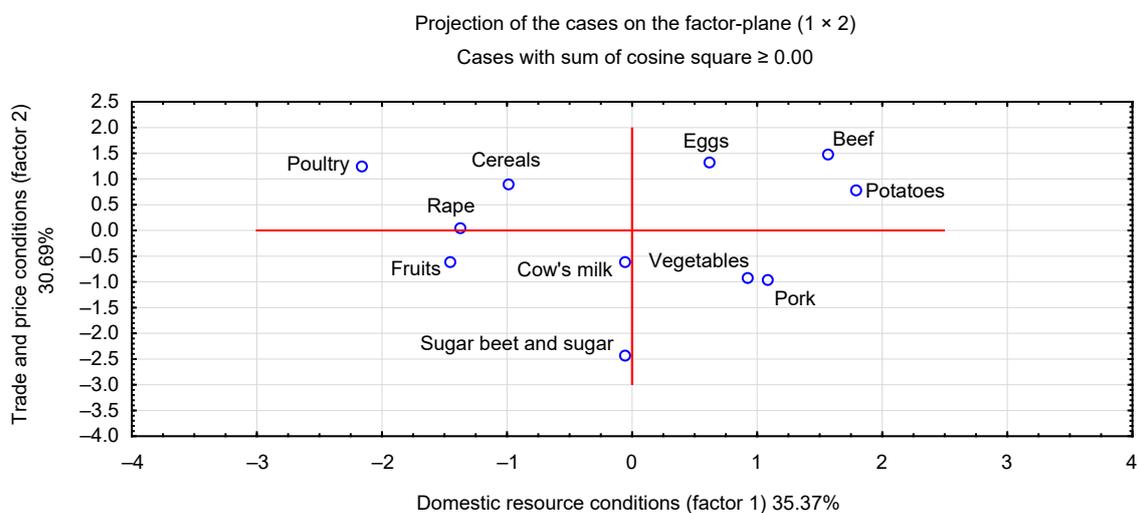


Fig. 1. Tested objects (agricultural product markets) mapped in the space of the first two principal components in the post-accession period
Source: same as in Table 1.

from a considerable increase in the volume of imports in 2009–2013, caused by the introduction of sugar quotas below the domestic production volume. In addition, the average purchase price of sugar beet was relatively low in the period concerned. The last aspect that made the sugar beet (sugar) market stand apart from other objects was a relatively small increase in the export volume. The distinctive feature of the second outlier (poultry market) was the lowest level of ‘domestic resource conditions.’ It means the largest standardized production volumes and domestic consumption levels were recorded on this market in the post-accession period (an inhibitory effect on the first component). Because of the value of the first factor, other three objects located within the shortest distance from the poultry market were the cereal, rape and fruit markets. The low value of the first component confirmed that all three of them, just like the poultry market, were characterized by a dynamic growth in the volume of domestic production and consumption.

As regards other objects, markets for cow’s milk, vegetables and pork were relatively close to each other. These sectors showed the least diversified values of the two principal components. On the other hand, the markets for potatoes, eggs and beef were similar to each other, especially because of a significant decrease in the volume of domestic consumption of products; as it

reached the smallest value for the eggs market, that market was closer to the ordinate axis than the other two. The potato market was the most distant from it because the decline in domestic demand in this market was accompanied by a decrease in production volumes. The above was not true for the eggs and beef market, where domestic production volumes increased over the study period due to the dynamic growth of exports. Hence, despite the similarity of these objects on the scatter plot, there was a significant difference between them as regards the causes behind the high value of the resource factor.

Although the choice of the number of clusters in the Ward’s method is arbitrary, there are some criteria and techniques that help making the decision. It is a common practice to use the agglomeration graph showing the distances between the clusters at the time of their creation. The cut-off line runs at the level where the graph becomes noticeably flat (Stanisz, 2007). During the agglomeration of individual groups in this study, the flattening is clearly visible, and extends to the ninth step. Therefore, considering the distances between successive links, the cut-off line should definitely be placed between the ninth and tenth step. On that basis, the objects were divided into three clusters. The clustering order and the distances between successive clusters are as

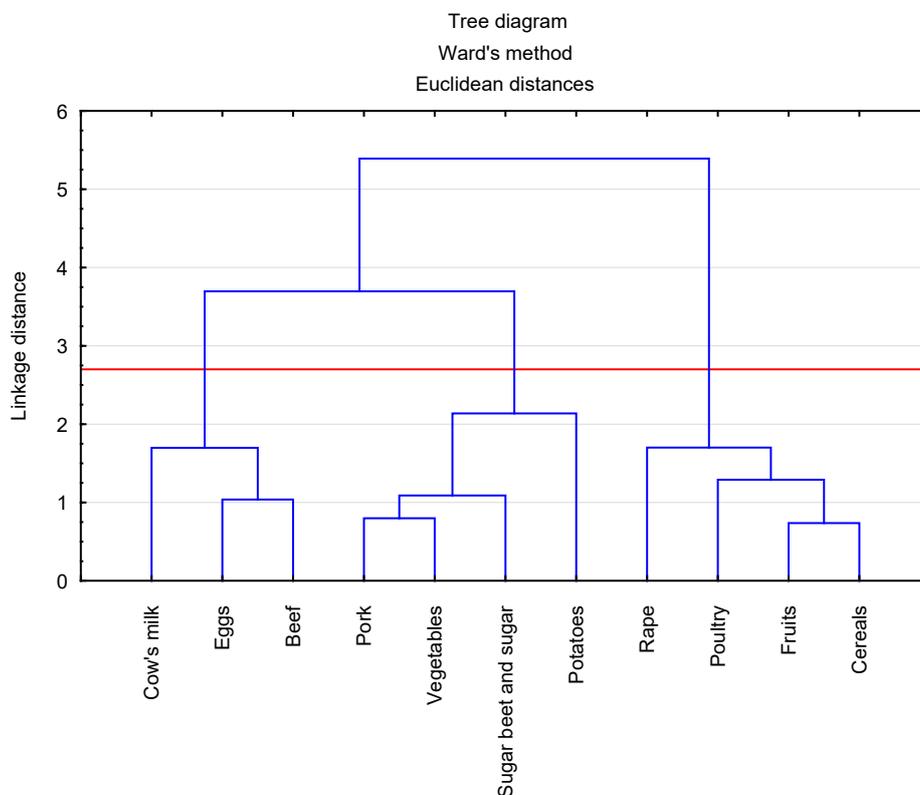


Fig. 2. Dendrogram of agricultural markets grouped into subsets by supply/demand conditions in the post-accession period
Source: same as in Table 1.

shown on the dendrogram (Fig. 2). The Ward's method resulted in the identification of the following clusters:

- 1 – cereals, rape, poultry, fruit,
- 2 – cow's milk, eggs, beef,
- 3 – potatoes, sugar beets (sugar), vegetables, pork.

The research procedure based on the k-means method resulted in the identification of the same clusters as in the case of the Ward's method. As shown in Figure 3, the greatest differences between the clusters were caused by variable x_1 (domestic production) and x_2 (domestic consumption). In turn, variable x_3 (export) had the smallest impact on the dispersion of objects. The average levels of features under consideration, as calculated for each group, revealed the differences between clusters.

The most favorable levels of variables under consideration were recorded in the first two groups. Note also the division of agglomerations by type of agricultural production. In the first group, except for poultry, there were three plant markets, while the second cluster

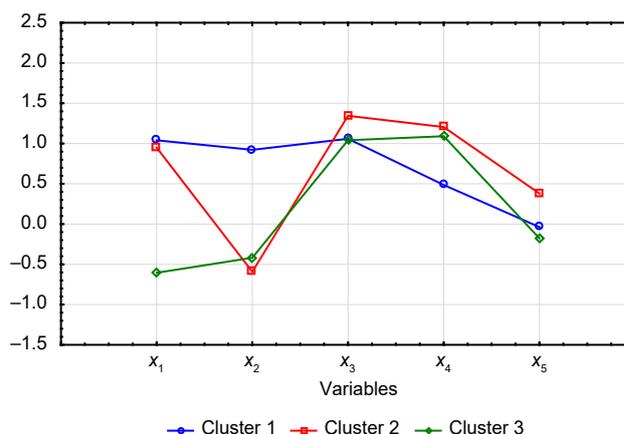


Fig. 3. Arithmetic mean of explanatory variables for the clustered markets for agricultural products in the post-accession period
Source: same as in Table 1.

consisted exclusively of animal markets. The last cluster, however, covered three markets for plant products and one for animal products.

CONCLUDING REMARKS

To conclude, the markets for agricultural products covered by this study can be divided into three aggregates which followed a similar development trend. The first group was composed of markets where both supply and demand demonstrated a consistent growth trend. This group includes cereal, rape, poultry and fruit markets (Czakowski, 2013). The characteristic shared by the objects in this agglomeration was a simultaneous increase in domestic supply and demand in the post-accession period (Czakowski, 2015; Czyżewski and Czakowski, 2016). In addition, these markets improved their international competitiveness, which was reflected by a consistently higher trade balance. On the other hand, because the (domestic and foreign) demand doubled, production levels were growing steadily in these markets, although no noticeable price increase was recorded in the post-accession period.

The second group covered markets demonstrating high levels of international competitiveness; in this case, the increase in production volume was primarily caused by the dynamic growth of exports in the post-accession period. This group included markets for cow's milk, eggs and beef. An additional stimulus for the development of these markets was a relatively favorable price formation in the post-accession period. However, the development of supply and demand on these markets was disturbed by a noticeable decline in domestic demand, as reflected by a significant surplus of domestic production. Nevertheless, it can be assumed that as long as domestic agricultural producers have a comparative advantage (mainly related to lower labor costs compared to EU countries), markets grouped in this cluster should continue enjoying favorable conditions.

The third group of markets extends to potatoes, sugar beets (sugar), vegetables and pork. Similarly to the first cluster, the supply/demand relationship remained stable in this group. However, this resulted from a decline in both quantities. Note that the reduction in potato consumption was caused by the decrease in both consumption and use of potatoes for processing and industrial purposes, while the core reason for the decrease in the quantity of pork produced was the poor

international competitiveness of this sector, resulting in a significant increase in imports to this market (Czakowski, 2012). The decline in vegetable and sugar beet yields was clearly smaller. In particular, the sugar beet market may expect the economic situation to get better soon, especially considering the abolition of sugar production limits.

REFERENCES

- Błażejczyk-Majka, L., Kala, R. (2005). Metody analizy skupień do charakterystyki użytków rolnych wybranych państw Unii Europejskiej [Cluster analysis methods for characterization agriculture area of selected members of the European Union]. *Roczn. Nauk. Stow. Ekon. Roln. Agrob.*, VII(5), 5–10 [in Polish].
- Cormack, R. (1971). A review of classification. *J. Royal Stat. Soc. Ser. A*, 134(3), 321–367.
- Czakowski, D. (2012). Dostosowania na rynku mięsa w Polsce przed i po akcesji do Unii Europejskiej [Adjustment for the meat market in Poland before and after accession to the European Union]. *Roczn. Ekon. KPSW*, 5, 61–74 [in Polish].
- Czakowski, D. (2013). Cereal Market before and after the Integration of Poland into the European Union. *Dynamics of Supply and Demand Correlations. Management* 17(1), 405–419.
- Czakowski, D. (2015). Zasobowe i efektywnościowe uwarunkowania produkcji drobiu i jaj w Polsce z perspektywy akcesji do Unii Europejskiej [Resources and effectiveness conditions of poultry and egg production in Poland from the perspective of accession to the European Union]. *Roczn. Nauk. Stow. Ekon. Roln. Agrob.*, 17(2), 20–25 [in Polish].
- Czyżewski, A. (1976). Miasta wielkopolskie w Polsce Ludowej: ekonomiczno-demograficzne podstawy rozwoju w okresie 1946–1970 [Wielkopolskie cities in People's Poland: economic and demographic foundations of development in the period 1946–1970]. Warszawa-Poznań: PWN [in Polish].
- Czyżewski, A., Guth, M. (2016). Zróżnicowanie produkcji mleka w makroregionach Unii Europejskiej z wyróżnieniem Polski [Diversification of milk production in macroregions of the European Union with the distinction of Poland]. Warszawa: Wyd. Nauk. PWN [in Polish].
- Czyżewski, A., Czakowski, D. (2016). Selected economic relationships on the fruit and vegetable market in Poland (1994–2013). *J. Agribus. Rural Dev.*, 4(42), 511–519.
- Czyżewski, A., Czakowski, D. (2017). Changes in Agricultural Markets in Poland following its Accession to the European Union. *Management*, 21(2), 138–150.

- Czyżewski, A., Strońska-Ziemann, J. (2014). Determinanty rozwoju obszarów wiejskich podregionu pilskiego [Factors determining the development of rural areas in pilski's sub-region]. *Roczn. Nauk. Stow. Ekon. Roln. Agrobiz.*, 16(3), 74–79 [in Polish].
- Fraley, C., Raftery, A. E. (1998). How many clusters? Which clustering method? Answers via model-based cluster analysis. *Comp. J.*, 41(8), 578–588.
- Grzelak, A. (2006). Wykorzystanie analizy skupień w badaniach struktur agrobiznesu na przykładzie powiązań gospodarstw rolnych z rynkiem [Using of cluster analysis in researches of structures of agribusiness on example of connections of farms with market]. *Zesz. Nauk. AR Wroc. Roln. LXXXVII (540)*, 179–185 [in Polish].
- GUS (1994–2013). *Rocznik Statystyczny Rzeczypospolitej Polskiej (dane za lata 1994–2013)* [Statistical yearbook of the Republic of Poland (data for the years 1994–2013)]. Warszawa: GUS [in Polish].
- GUS (1995–2014). *Skup i ceny produktów rolnych (dane za lata 1994–2013)* [Purchase and prices of agricultural products (data for the years 1994–2013)]. Warszawa: GUS [in Polish].
- Jolliffe, I. (2002). *Principal component analysis*. Second Edition. Springer.
- Kaiser, H. F. (1960). The Application of Electronic Computers to Factor Analysis. *Educ. Psych. Measur.*, XX(1), 141–151.
- Krzanowski, W. (2000). *Principles of multivariate analysis. Analysis: A user's perspective*. Oxford: University Press.
- MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations, *Proc. of the fifth Berkeley symp. on math. statist. and prob.*, 1(14), 281–297.
- Marek, T. (1989). *Analiza skupień w badaniach empirycznych: metody SAHN* [Cluster analysis in empirical research: SAHN methods]. Warszawa: PWN [in Polish].
- Mądry, W., Gozdowski, D., Roszkowska-Mądra, B., Hryniewski, R. (2011). *Typologia systemów produkcji rolnej: koncepcja, metodologia i zastosowanie* [Typology of farming systems: concept, methodology and use]. *Fragm. Agron.*, 28(3), 70–81.
- Panek, T., Zwierzchowski, J. (2013). *Statystyczne metody wielowymiarowej analizy porównawczej: teoria i zastosowania* [Statistical methods of multivariate comparative analysis: theory and applications]. Warszawa: SGH [in Polish].
- Pietrzykowski, R., Kobus, P. (2006). Zastosowanie modyfikacji metody k-średnich w analizie portfelowej [Application of modify k-means method to portfolio analysis]. *Ekon. Org. Gosp. Żywn.*, 60, 301–308 [in Polish].
- Pietrzykowski, R., Kobus, P. (2008). *Wielowymiarowe metody statystyczne w analizie wyników ekonomiczno-produkcyjnych gospodarstw rolnych wybranych państw Unii Europejskiej* [Multivariate Statistical Methods in Analysis of Production and Economic Results of Agricultural Holdings in the EU Countries]. ZN SGGW w Warszawie. *Probl. Roln. Świat.*, 4(19), 371–378 [in Polish].
- Poczta, W., Pawlak, K. (2008). *Typologia wzorców konsumpcji podstawowych artykułów żywnościowych w krajach Unii Europejskiej* [Typology of Food Consumption Patterns in the European Union Countries]. *Roczn. Nauk. Stow. Ekon. Roln. Agrob.*, VII (8), 195–203 [in Polish].
- Stanisz, A. (2007). *Przystępny kurs statystyki z zastosowaniem STATISTICA PL na przykładach z medycyny. Analizy wielowymiarowe, Tom 3* [Affordable statistic course using STATISTICA PL Vol. 3]. Kraków: StatSoft [in Polish].